SESSION

CLOUD COMPUTING + WEB SERVICES + WEB APPLICATIONS + SYSTEM INFRASTRUCTURES AND ENHANCED CACHE METHODS

Chair(s)

TBA

NbQ-CLOCK: A Non-blocking Queue-based CLOCK Algorithm for Web-Object Caching

Gage Eads[1], **Juan A. Colmenares**[2]
[1]EECS Department, University of California, Berkeley, CA, USA
[2]Computer Science Laboratory, Samsung Research America – Silicon Valley, CA, USA
geads@eecs.berkeley.edu, juan.col@samsung.com

Abstract— *Major Internet-based service providers rely on high-throughput web-object caches to serve millions of daily accesses to frequently viewed web content. A web-object cache's ability to reduce user access time is dependent on its replacement algorithm and the cache hit rate it yields. In this paper, we present NbQ-CLOCK, a simple and effective lock-free variant of the Generalized CLOCK algorithm that is particularly suited for web-object caching. NbQ-CLOCK is based on an unbounded non-blocking queue with no internal dynamic memory management, instead of the traditional circular buffer. Our solution benefits from Generalized CLOCK's low-latency updates and high hit rates, and its non-blocking implementation makes it scalable with only 10 bytes per-object space overhead. We demonstrate that NbQ-CLOCK offers better throughput than other competing algorithms, and its fast* `update` *operation scales well with the number of threads. We also show that for our in-memory key-value store prototype, NbQ-CLOCK provides an overall throughput improvement of as much as 9.20% over the best of the other algorithms.*

Keywords: Replacement algorithm, key-value cache, CLOCK, non-blocking, scalability

1. Introduction

Minimizing the service response time experienced by users is very important for global-scale Internet-based service providers, such as Amazon, Facebook, and Samsung. One way to reduce service response times is with web-object caching, in which recently or frequently accessed remote data is cached locally to avoid slow remote requests when possible.

Web-object caches are often implemented as key-value stores. In general, key-value stores provide access to unstructured data through read and write operations, where each unique key maps to one data object. Popular key-value stores used for web-object caching operate purely *in memory*; a prime example is Memcached [1] (summarized in Section 2). In key-value stores of this type, frequently accessed data items are stored in cache servers' volatile RAM to achieve low round-trip access times. Only object requests that miss in the cache are routed to the slower, non-volatile storage layer; in this way, the key-value store mitigates the load on the storage layer.

G. Eads conducted this work during his internship at Samsung Research America – Silicon Valley

A recent analysis [2] finds that Memcached on Linux spends 83% of its processing time in the kernel, primarily in the network layer. However, recent work such as KV-Cache [3] removes this bottleneck by using a lightweight protocol stack together with a zero copy design. Furthermore, non-blocking hash tables can eliminate contention accessing key-value pairs. Thus, we anticipate the bottleneck in high-performance key-value stores shifting to the replacement algorithm, the focus of this paper.

The replacement algorithm is responsible for deciding which item to evict when the cache is full and there are new incoming items. The approach used to evict items in a web-object cache can dramatically impact the performance of the Internet-based services the cache assists, and that influence on service performance is characterized by the *cache hit rate*.

In the context of web-object caching, a replacement algorithm's API consists of four operations: `insert` adds a new item to the replacement data structure, `delete` removes an item from the replacement data structure, `update` notifies the replacement algorithm of an access of an existing item, and `evict` selects one or more cached items for eviction.

A recent study of Facebook's Memcached traces [4] shows that large-scale key-value store workloads are read-heavy, and key and value sizes vary. Hence, a replacement algorithm for an in-memory key-value store should support variably-sized objects and allow for low-latency, scalable updates and evictions, as well as high cache hit rates.

The most common replacement algorithms are the *Least Recently Used* (LRU) algorithm and its derivatives. LRU, as its name implies, evicts the least recently accessed item. Its derivatives (*e.g.,* pseudo-LRU and CLOCK) trade-off hit rate in favor of lower space complexity or implementation cost.

CLOCK [5] is a well-known memory-page replacement algorithm. It maintains a circular buffer of reference bits, one for each memory page, and a buffer pointer ("clock hand"). When a page is referenced, its reference bit is set to indicate a recent access. To evict a page, the clock hand sweeps through the buffer and resets each non-zero bit it encounters, until it finds an unset bit; then, the corresponding page is evicted. Unfortunately, CLOCK's hit rate can suffer because with a single bit reference it cannot differentiate access frequency from access recency. To solve this problem, Generalized CLOCK [6] replaces each reference bit with a counter.

While their fast update path is ideal for read-mostly web-object caches, CLOCK and its variants assume a *fixed* number of cache entries. For instance, CAR (CLOCK with Adaptive Replacement) [7] assumes the cache has a fixed size c, and CLOCK-pro [8] depends on a fixed total memory size m, and both algorithms operate on uniformly-sized pages. WSClock [9] is based on a fixed circular list. The fixed-size assumption is *not* the case for web-object caches, and this limitation renders CLOCK and existing variants impractical in the web-object caching domain.

In this paper, we present **N**on-**b**locking **Q**ueue-based **CLOCK** (NbQ-CLOCK). It is a lock-free variant of the Generalized CLOCK replacement algorithm suited to caches containing variable number of items with different sizes. To the best of our knowledge, no prior work in literature has evaluated a similar CLOCK variant that could effectively handle such dynamically-sized caches, which are essential to web-object caching. NbQ-CLOCK has been implemented in KV-Cache [3], a high-performance in-memory key-value cache conforming to the Memcache protocol, with encouraging preliminary results.

In Section 5, we evaluate NbQ-CLOCK against other replacement algorithms applicable to Memcached, including Bag-LRU [10]. We focus our evaluation on Memcached [1] because it is widely deployed and has recently received considerable attention [10], [11], [12]. We demonstrate that NbQ-CLOCK's low update latency scales well with the thread count, and that it exhibits better throughput scaling than the other algorithms. Moreover, when used in our in-memory key-value store prototype, NbQ-CLOCK offers an improvement on the system's throughput of as much as 9.20% over the best of the other replacement algorithms. We also show that NbQ-CLOCK's cache hit rates are at least as good as those of the other replacement algorithms.

2. An Overview of Memcached

Memcached [1], [12], [13] is a widely deployed web-object caching solution. It is typically deployed in a "side-cache" configuration, in which end users, via client devices, send requests to the front-end web servers. The front-end servers then attempt to resolve each end-user request from one or more local Memcached servers by, in turn, sending `GET` requests to them. If a cache miss occurs, the front-end server handling the end-user request forwards it to the back-end database servers that carry out the computation and IO operations to produce the result. On receiving the result, the front-end server both sends the answer to the client and updates the cache by issuing a `SET` request to the appropriate Memcached server.

An analysis of Facebook's Memcached traces over a period of several days [4] revealed the following about large-scale key-value store workloads. They are read-heavy, with a `GET`/`SET` ratio of 30:1, and request sizes are seen as small as 2 bytes and as large as 1 Mbytes.

3. Replacement Algorithms for Memcached

In this section, we describe two existing replacement algorithms used in Memcached [1]: the algorithm shipped with its stock version, and Bag-LRU [10]. In addition, we present *Static CLOCK*, a conceivable, but ineffective extension to CLOCK that statically over-provisions for the largest number of items that an instance of an in-memory key-value store (*e.g.,* Memcached) can possibly accommodate in order to support variably-sized objects.

3.1 Memcached LRU

The replacement algorithm in the stock version of Memcached is a per-slab class LRU algorithm. Every item in Memcached has a corresponding item descriptor, which contains a pointer for the LRU list. This algorithm uses a doubly-linked list to support arbitrary item removals, which is important for `DELETE` requests and expired items. A global cache lock protects the LRU list from concurrent modifications, and fine-grained locks protect the hash table. Lock contention on the global lock greatly impairs intra-node scalability, though this was not a concern at the time of Memcached's initial development. Instead, the developers sought a simple mechanism to ensure thread-safety.[1] This synchronization solution is clearly unscalable, and there have been many attempts at solving it.

3.2 Bag-LRU

Memcached's poor scalability within the node motivated the development of the Bag-LRU replacement strategy [10]. Bag-LRU is an LRU approximation designed to mitigate lock contention. Its data structure comprises a list of multiple timestamp-ordered "bags", each containing a pointer to the head of a singly-linked list of items. Bag-LRU keeps track of the two newest bags (needed for updates) and the oldest bag (needed for evictions).

Bag-LRU's `update` operation consists of writing the newest bag's address to the item's back pointer and updating the item's timestamp. The lock-free `insert` operation places the item in the newest bag's list. To do so, a worker thread repeatedly attempts to append the item to the tail of the list using the atomic compare-and-swap (CAS) operation.

[1] Brad Fitzpatrick, Memcached's original developer, built it with a "scale out, not up" philosophy [14], at a time when multi-core chips were just entering the market. Scale-up, however, is important for Memcached deployments to rapidly serve web-objects stored on the same node in parallel.

If the CAS fails, the worker thread traverses the bag's list until it finds a NULL next pointer and retries.

An eviction requires grabbing a global eviction lock to determine the oldest bag with items in it, locking that bag, and choosing the first available item to evict. Bag-LRU requires locks for evictions and deletes to prevent the cleaner thread from simultaneously removing items, which can result in a corrupted list.

3.3 Static CLOCK

There are two ways to extend CLOCK in order to support dynamically-sized objects. These are:

- Statically pre-allocate a bit-buffer that is large enough to support the worst-case (*i.e.,* most) number of items. We refer to it as *Static CLOCK*.
- Replace the underlying data structure with a list. This is the alternative we pursue in this paper and our solution is described in Section 4.

For completeness, we evaluate Static CLOCK, and show that the list-based approach is much more effective for web-object caching.

4. NbQ-CLOCK

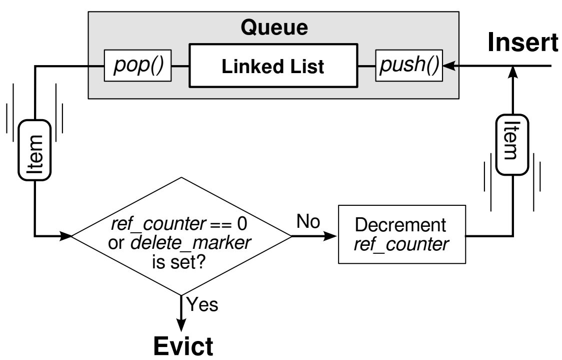


Fig. 1: Process flow of queue-based `evict`, `delete`, and `insert` operations. The data items are the moving elements, as opposed to the clock hand in the traditional CLOCK.

Non-**b**locking **Q**ueue-based **CLOCK** (NbQ-CLOCK) is a lock-free variant of the Generalized CLOCK replacement algorithm [6]. The primary difference between NbQ-CLOCK and previous CLOCK variants is that it circulates the cached items through the eviction logic (see Figure 1), instead of iterating over the items (by moving the clock hand). The efficient and scalable circulation of items is enabled by the use of a non-blocking concurrent queue, as opposed to the traditional statically allocated circular buffer in prior CLOCK variants. Moreover, the use of an unbounded queue allows NbQ-CLOCK to handle a dynamically-sized cache (containing variable number of items with different sizes), which is key for web-object caching (*e.g.,* Memcached [1]).

NbQ-CLOCK's non-blocking queue is based on a singly-linked list, and its algorithmic details are presented next in Section 4.1. NbQ-CLOCK stores bookkeeping information in each linked-list node; a node includes:

- a reference counter,[2]
- an atomic delete marker, and
- a pointer to the next node.

In addition, each linked-list node contains a pointer to a unique data item (*e.g.,* a key-value pair). We also refer to these nodes as *item descriptors*.

NbQ-CLOCK maintains the traditional CLOCK interface with the following operations:

- `insert` pushes an allocated list node into the queue. The node's reference counter is initialized to zero and the delete marker to `false`.
- `update` increments the node's reference counter.
- `delete` removes an item from the queue. The non-blocking queue can only remove items from the head of the queue, but web-object caches must support the ability to delete arbitrary objects. To support arbitrary object deletions, we include a "delete marker" in each item descriptor. This binary flag is set atomically and indicates whether the corresponding item is deleted from the cache. During the clock sweep operation, the worker thread frees the memory of any items whose delete marker is set.
- `evict` pops the head of the queue, and if the node's reference counter is zero or its delete marker is set, evicts the item. If the item is not suitable for eviction, its reference counter is decremented and it is recycled back into the queue.

Figure 1 depicts NbQ-CLOCK's `insert`, `delete`, and `evict` operations. For more details on the design of NbQ-CLOCK, please refer to [15].

4.1 Underlying Non-blocking Queue

NbQ-CLOCK is operates on top of a non-blocking concurrent queue. Hence, the clock hand is not an explicitly maintained variable, but instead is implicitly represented by the head of the queue.

For our non-blocking queue we use Michael and Scott's algorithm [16], but we have made two key optimizations. First, our queue performs no memory allocation in its `init`, `push`, or `pop` operations. Instead, the caller function is responsible for allocating and freeing nodes. Second, NbQ-CLOCK's `push` and `pop` methods operate on nodes, not the data items themselves, so that one can re-insert a popped node at the tail with no memory management overhead. These optimizations primarily benefit NbQ-CLOCK's `evict` operation.

Michael and Scott's algorithm consists of a singly-linked list of nodes, a tail pointer, and a head pointer, where the

[2] or a reference bit for a non-blocking variant of the original CLOCK [5].

head always points to a dummy node at the front of the list. Their algorithm uses a simple form of snapshotting, in which the pointer values are re-checked before and during the CAS operations, to obtain consistent pointer values.

One manner in which NbQ-CLOCK's queue and Michael and Scott's algorithm differ is that their queue performs a memory allocation on each `push` and a de-allocation on each `pop`. Then, if a clock sweep operation inspects n items before finding one suitable for eviction, it pops n list nodes and pushes $(n-1)$ list nodes – resulting in n memory frees and $(n-1)$ memory allocations. On the contrary, our non-blocking queue takes the memory management logic out of `push` and `pop` so that every `evict` incurs the minimum amount of memory management overhead: a single de-allocation. Further, most memory allocators use internal locking, in which case the queue is not fully non-blocking. We avoid this problem by taking the allocators out of the queue.

We also remove the modification counters in Michael and Scott's algorithm from our queue. The modification counters protect against the ABA problem,[3] which can corrupt the queue's linked list. However, for sufficiently large queues, the likelihood that other threads can cycle through the queue in such a short time is effectively zero. Since key-value stores typically cache millions of items, we choose not to include modification counters in the NbQ-CLOCK queue.

5. Comparative Evaluation

In this section, we evaluate NbQ-CLOCK against the replacement algorithms Memcached LRU, Bag-LRU, and Static CLOCK, presented in Section 3. We focus on *update latency scalability*, *cache hit rate*, and *throughput*, which are performance metrics important to high-throughput, low-latency, read-mostly web-object caches.

The version of NbQ-CLOCK used in our experiments implements the Generalized CLOCK with 8-bit reference counters. The Static CLOCK, on the other hand, uses an array of packed reference bits.

5.1 Experimental Setup

The test software platform consists of two C++ applications: an *in-memory key-value store prototype*, which is Memcached-protocol conformant, and a *client traffic generator*, called KVTG. The two applications run in Linux 3.2.0 and are connected through high-performance shared-memory channels [17].

The key-value store prototype primarily comprises a hash table and an item-replacement algorithm, plus the logic necessary to receive and parse client requests and generate and transmit responses. Each worker thread in the key-value store receives requests on a receive (RX) channel, performs the hash-table and replacement operations necessary to satisfy the requests, and transmits the results across a transmit (TX) channel.

The key-value store prototype, like Memcached [1], allows us to set the cache size, which is the memory limit for keys, values, and item descriptors (including replacement-algorithm bookkeeping). Also similar to Memcached, the prototype uses 7 slab allocators with sizes ranging in powers of two from 64 B to 4 KB. While per-slab class replacement logic is necessary in a deployment scenario,[4] we restrict our focus to a single replacement instance by choosing object sizes that use a single slab allocator.

Our experimental platform is a quad-socket server containing the Intel E5-4640 2.4GHz CPU (8 cores, 16 hardware threads per socket) with 20 MB of last-level cache, and 128-GB DRAM overall. To minimize performance variability across test runs, we disable Turbo Boost, redirect interrupts to unused cores, fix each CPU's frequency to 2.4 GHz, and affinitize software threads to cores isolated from the Linux scheduler.

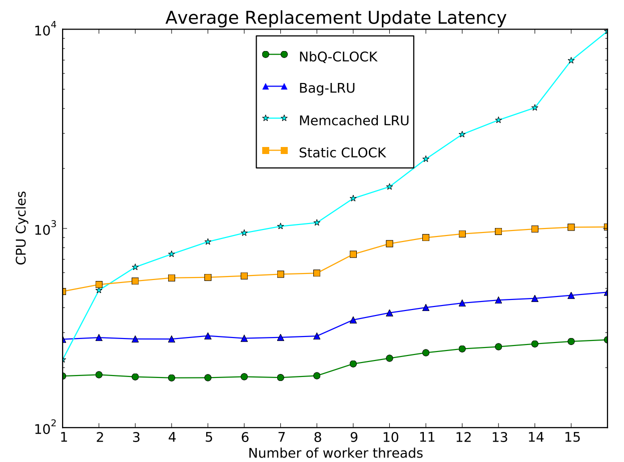


Fig. 2: Scalability of replacement algorithms' `update` operation. Each data point is the average `update` latency over 10 million `GET` requests.

5.2 Update Latency Scaling

Update latency is the time to update a cached item's entry in the replacement data structure in the event of a cache hit. It is crucial for server performance that this process – the common case in a read-heavy workload – scales well.

In this experiment, we measure the average latency for the `update` operation of the replacement algorithms under evaluation with the worker thread count varying from 1 to 16. The results are shown in Figure 2, with the y-axis (CPU cycles) plotted on a logarithmic scale. Each data point in

[3] The ABA problem can occur when, between reading a shared value of A and performing a CAS on it, another thread changes the A to a B and then back to n A. In this case, the CAS may succeed when it should not.

[4] When a request causes an eviction, the evicted item must occupy enough dynamic memory to satisfy that request. To find an appropriately-sized item quickly, each slab class must have its own replacement logic.

the figure was calculated over 10 million `GET` requests, with key frequency determined by a power-law probability distribution. NbQ-CLOCK's update latency scales significantly better than Memcached LRU. Heavy contention for the global lock seriously hinders LRU's scalability, resulting in a 44.6x increase in mean update latency from 1 to 16 threads, while NbQ-CLOCK's mean update latency increases by 2.11x in that range.

Interestingly, Static CLOCK's `update` performs worse than NbQ-CLOCK despite both being CLOCK variants. The main reason is the read-modify-write loop with CAS in the `update` operation of our Static-CLOCK implementation. We observe that a single atomic CAS operation takes 483 cycles on average on the test platform for a single thread (see Figure 2), but the likelihood of CAS failing and the loop repeating grows with the number of threads, resulting in the observed update-latency increase. NbQ-CLOCK's non-blocking `update` operation, requiring a single increment operation, scales much better.

The Bag-LRU solution [10] scales well, but its average update latency is between 1.5x-1.73x that of NbQ-CLOCK. This is because, besides writing the back pointer, the `update` operation must also check if the newest bag is full and, if so, atomically update the global newest bag pointer.

The performance dip between 8 and 9 worker threads occurs in all algorithms and appears to be a memory-system artifact resulting from the 9th thread running on a second socket.

5.3 Cache Hit Rate

Another metric fundamental to a web-object cache is hit rate. Besides having lower-latency `update` operations, this experiment shows that NbQ-CLOCK is comparable to and never worse than the alternative algorithms in terms of hit rate.

We set a memory limit of 1 GB for slab allocators for keys, values, item descriptors, and replacement algorithm overhead. The key space is modeled by a standard normal distribution of 30 million items and the requests are 70% `GET`s and 30% `ADD`s. KVTG issues 15 million `SET` requests during the "warm up" phase, in which no data is collected. In the subsequent phase, 15 million requests are sent according to the key and request-type distributions and hit rate data is collected. We evaluate a range of object sizes (key plus value) from 64 B to 4 KB.

Table 1 shows that NbQ-CLOCK's cache hit rates exceed the next best algorithm by as much as 1.40% (for 4 KB objects). NbQ-CLOCK's hit rate improvement over Bag-LRU's is more significant in the context of real workloads of web-object cache systems. For instance, considering the workload characterization of live Memcached traffic reported in [4], an additional 1.40% hits of 4.897 billion requests in a day amounts to an additional 68.6 million cache hits per day.

Object Size (bytes)	NbQ-CLOCK	Bag-LRU	Static CLOCK	LRU
	Number of items stored (millions)			
64	**5.263**	5.113	4.503	5.089
128	**4.006**	3.919	3.336	3.905
256	**2.711**	2.671	2.198	2.664
512	**1.647**	1.632	1.306	1.629
1024	**0.922**	0.918	0.721	0.916
2048	**0.491**	0.489	0.380	0.489
4096	**0.253**	**0.253**	0.196	**0.253**
	Hit rate			
64	**82.81%**	82.34%	81.32%	82.29%
128	**80.86%**	80.76%	80.34%	80.76%
256	**78.23%**	78.13%	77.14%	78.13%
512	**75.26%**	**75.26%**	74.06%	75.25%
1024	**72.20%**	72.05%	70.90%	72.07%
2048	**68.92%**	68.32%	67.29%	68.32%
4096	**65.60%**	64.20%	63.32%	64.19%

Table 1: Number of stored data items and hit rate for a 1 GB cache, with a standard normal key distribution, across various object sizes.

Furthermore, this workload characterization was generated from five server pools within one of many datacenters; the impact of an improved replacement algorithm compounds in a global Memcached deployment.

NbQ-CLOCK's higher hit rates for small objects can be attributed to its superior space efficiency. Compared to Bag-LRU, NbQ-CLOCK has six fewer bytes per item due to its use of one pointer for its singly-, not doubly-, linked list, plus the single-byte reference counter and single-byte delete marker. For smaller object sizes, NbQ-CLOCK's bookkeeping space advantage is more significant relative to the object size, allowing it to store more objects than the alternatives – 150 thousand more than Bag-LRU for 64 B objects, for instance.

Static CLOCK has poor space efficiency, despite our effort to improve this aspect in our implementation. The reason is that it requires a statically allocated reference-bit array and item pointer array for every possible item in the cache. In this experiment, 249.67 MB (25% of the memory limit) is dedicated to Static CLOCK's arrays, and the hit rate suffers as a result.

For larger object sizes, the space overhead advantage of NbQ-CLOCK is insignificant with respect to the object size – NbQ-CLOCK, Bag-LRU, and LRU can store nearly the same number of items for 2 KB and higher. In the case of 4 KB objects, NbQ-CLOCK, Bag-LRU, and LRU can only store a small fraction (0.84%) of the 30 million items (due to the 1 GB memory limit), so the eviction algorithm's intelligence is the main factor affecting the hit rate.

The reason for NbQ-CLOCK's hit-rate advantage for 1 KB and larger objects is that NbQ-CLOCK, as a variant of Generalized CLOCK, considers not only access recency, but also access frequency. As a result, frequently accessed items

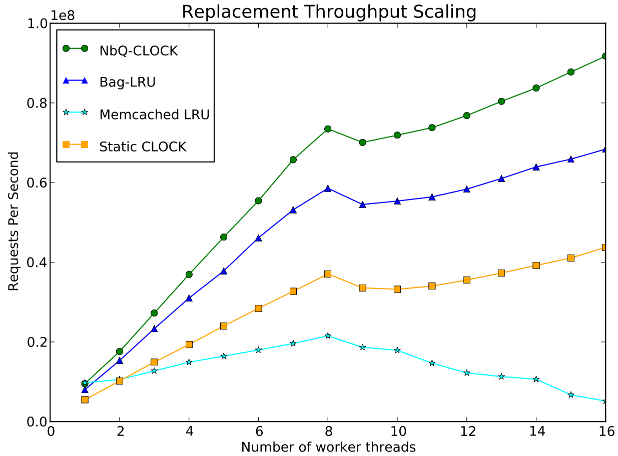


Fig. 3: Throughput scaling of the replacement algorithms when the in-memory key-value store imposes no additional bottlenecks.

tend to persist longer in a cache with Generalized CLOCK than in one with LRU. This is particularly important when the number of key-value pairs the cache can store is small compared to the total set of pairs, and key appearance is governed by a power-law distribution (as in this experiment). For this distribution, a small fraction of the keys appear much more frequently than the rest, and NbQ-CLOCK keeps the key-value pairs in that fraction longer than the LRU variants do.

5.4 Throughput Scaling

In the following two experiments, we measure the scalability of the replacement algorithms in isolation and as part of a key-value store. In the first experiment, we evaluate the performance of NbQ-CLOCK, Memcached LRU, Bag-LRU, and Static CLOCK in the best case – *i.e.,* when the rest of the cache imposes no bottlenecks. This allows one to accurately measure the scalability (or lack thereof) of each replacement algorithm in the cache. To remove all bottlenecks, we replace the hash table with a pre-allocated array of keys and item descriptors and remove any dynamic memory management. By designing the benchmark in this way, we replicate the behavior of the hash table without any of its scalability bottlenecks.

For this experiment, we define throughput as the time spent in the replacement algorithm operations divided by the number of requests (10 million). We use 70% `GET` and 30% `ADD` operations. We model the key-appearance frequency with a power-law distribution for 10 million unique keys, and use an object size of 128 B. We fill the cache during an initialization phase before running the experiment.

The results are shown in Figure 3. Each cache stores between 33% and 40% of the 10 million objects, depending on the replacement algorithm. By storing at least 1/3 of the keys whose appearance is modeled by a power-law distribution, 91% of all `GET`s are successful. On the other hand, successful `ADD`s - wherein the key does not already exist - occur the other 9% of the time. NbQ-CLOCK, with its low-latency updates (276 cycles average for 16 threads), outperforms the other replacement algorithms. The other algorithms perform as one would expect from Figure 2: Bag-LRU performs better than Static CLOCK, and Memcached LRU scales poorly, particularly after crossing the socket boundary. As with the update latency scaling experiment in Section 5.2, the performance dip between 8 and 9 worker threads appears to be a memory system artifact resulting from the 9th thread running on a second socket.

While hit rate and update latency scalability give important insight into the performance of a replacement algorithm, those pieces in isolation can only tell part of the story. To capture the effect of a replacement algorithm on the performance of an in-memory key-value store, one must measure the entire system's throughput. Unlike in the previous benchmark, in the following benchmark the key-value store prototype *is not* fully optimized for scalability; there is lock contention in the hash table and memory management operations. Thus NbQ-CLOCK's inherent advantages are hindered by the rest of the cache.

For the second scaling experiment, we set a memory limit of 1 GB for slab allocators for keys, values, item descriptors, and replacement algorithm overhead, and use an object size of 128 bytes. The key-appearance frequency is modeled by a power-law distribution of 10 million keys and the requests are 70% `GET`s and 30% `ADD`s. KVTG issues 40 million `SET` requests during the "warm up" phase to ensure the cache is well populated; in this phase no data is collected. In the subsequent phase, 20 million requests are sent to each worker thread according to the key-appearance and request-type distributions. We evaluate each replacement algorithm for up to 6 worker threads. For up to 4 worker threads the key-value store prototype uses the same thread configuration as that in the latency scaling experiment (Section 5.2). However, for 5 and 6 worker threads, we pair a KVTG's transmit thread and a key-value store's worker thread on sibling hardware threads.

The results are shown in Figure 4. As expected, there is much less differentiation in throughput between the four replacement algorithms with the key-value store prototype. For one and two worker threads, and again with five and six worker threads, there is little difference between the algorithms. As opposed to the previous experiment, the latency for a single request is a function of the replacement algorithm *and* the rest of the cache (*i.e.,* the hash table and memory management). As the number of worker threads increases and the performance difference in the replacement algorithm grows, the throughput difference becomes more pronounced. With four worker threads, NbQ-CLOCK's throughput exceeds the next best by 9.20%. Because the rest of the cache is the limiting factor for five or more worker

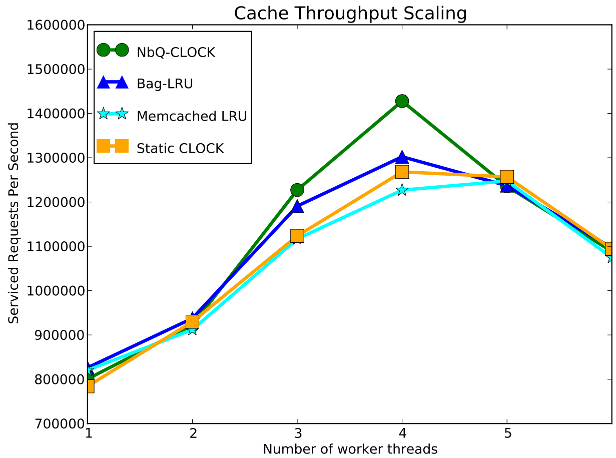


Fig. 4: Throughput scaling of a realistic in-memory key-value store with different replacement algorithms.

threads, we do not show results beyond six threads.

6. Related Work

Nb-GCLOCK [18] is a non-blocking variant of the Generalized CLOCK algorithm intended for memory-page replacement in operating systems. Nb-GCLOCK assumes a fixed (typically 4KB) cache-object size, and as such is based on a statically allocated circular buffer. This assumption typically does not hold for in-memory key-value stores (*e.g.,* Memcached), where keys and values have *variable size*, which makes Nb-GCLOCK unsuitable for most web-object caching scenarios.

MemC3 [11] is a Memcached implementation that uses CLOCK for its replacement algorithm. However, MemC3 makes the (generally incorrect) assumption that the object size is fixed. Further, MemC3 does not consider the possibility of slab class re-balancing, which requires either pre-allocating a large enough CLOCK buffer for the worst case (*i.e.,* re-balancing all memory to a given slab) for every slab, or dynamic CLOCK buffer resizing. However, maintaining correctness in the presence of dynamic CLOCK buffer resizing is impossible without introducing locks, and statically pre-allocating the buffer introduces an unwieldy space overhead (discussed further in Section 4).

7. Conclusion

This paper presents NbQ-CLOCK, a replacement algorithm for web-object caches. Its performance exceeds state-of-the-art algorithms like Bag-LRU in terms of overall system throughput and number of items stored, while offering hit rates at least as good as those of the alternative algorithms. Furthermore, NbQ-CLOCK's simplicity is beneficial when developing a multi-threaded key-value store, as it requires less debugging and testing effort than more complicated alternatives.

NbQ-CLOCK has been implemented in KV-Cache [3], a high-performance in-memory key-value cache conforming to the Memcache protocol. Our preliminary experimental results show that, when servicing traffic consisting of 70% `GET`s and 30% `SET`s, KV-Cache experiences negligible degradation in total system throughput, whereas the performance of Intel's Bag-LRU Memcached [10] is severely affected by the presence of `SET` requests. Besides storing and replacing data items, `SET` requests trigger KV-Cache's eviction logic once the amount of memory used to store key-value pairs exceeds a predetermined threshold. Thus, these results are an encouraging indication of NbQ-CLOCK's effectiveness as part of this full system.

One future direction for this work is to adapt higher-performance CLOCK variants (*e.g.,* [7], [8], [9]) to web-object caching with the ideas presented in this paper. CLOCK has well-documented deficiencies that these variants overcome. For instance, they can cope with scans, self-tune to a given workload, better measure access frequency, and in general outperform CLOCK. Additionally, we also believe that NbQ-CLOCK could be used in other domains, especially when the cache size and object size vary. We plan to explore these opportunities in the future.

References

[1] "Memcached," http://www.memcached.org.

[2] J. Leverich and C. Kozyrakis, "Reconciling high server utilization and sub-millisecond quality-of-service," in *Proceedings of the 9th European Conference on Computer Systems (Eurosys'14)*, April 2014.

[3] D. Waddington, J. Colmenares, J. Kuang, and F. Song, "KV-Cache: A scalable high-performance web-object cache for manycore," in *Proceedings of the 6th IEEE/ACM International Conference on Utility and Cloud Computing (UCC'13)*, December 2013, pp. 123–130.

[4] B. Atikoglu, Y. Xu, E. Frachtenberg, S. Jiang, and M. Paleczny, "Workload analysis of a large-scale key-value store," *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1, pp. 53–64, June 2012.

[5] F. J. Corbató, "A paging experiment with the Multics system," in *In Honor of P. M. Morse*. MIT Press, 1969, pp. 217–228.

[6] A. J. Smith, "Sequentiality and prefetching in database systems," *ACM Transactions on Database Systems*, vol. 3, no. 3, pp. 223–247, September 1978.

[7] S. Bansal and D. S. Modha, "CAR: Clock with adaptive replacement," in *Proceedings of the 3rd USENIX Conference on File and Storage Technologies (FAST'04)*, March 2004, pp. 187–200.

[8] S. Jiang, F. Chen, and X. Zhang, "CLOCK-Pro: An effective improvement of the CLOCK replacement," in *Proceedings of the 2005 USENIX Annual Technical Conference (ATC'05)*, April 2005, pp. 323–336.

[9] R. W. Carr and J. L. Hennessy, "WSCLOCK: A simple and effective algorithm for virtual memory management," in *Proceedings of the 8th ACM Symposium on Operating Systems Principles (SOSP'81)*, 1981, pp. 87–95.

[10] A. Wiggins and J. Langston, "Enhancing the scalability of Memcached," Intel Corporation, Tech. Rep., May 2012. [Online]. Available: http://software.intel.com/en-us/articles/enhancing-the-scalability-of-memcached

[11] B. Fan, D. G. Andersen, and M. Kaminsky, "MemC3: Compact and concurrent MemCache with dumber caching and smarter hashing," in *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (NSDI'13)*, April 2013, pp. 371–384.

[12] R. Nishtala, H. Fugal, S. Grimm, M. Kwiatkowski, H. Lee, H. C. Li, R. McElroy, M. Paleczny, D. Peek, P. Saab, D. Stafford, T. Tung, and V. Venkataramani, "Scaling Memcache at Facebook," in *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (NSDI'13)*, April 2013, pp. 385–398.

[13] C. Aniszczyk, "Caching with Twemcache," https://blog.twitter.com/2012/caching-twemcache, 2012.

[14] B. Fitzpatrick, "Distributed caching with Memcached," *Linux Journal*, vol. 2004, no. 124, pp. 5–, August 2004.

[15] G. Eads, "NbQ-CLOCK: A non-blocking queue-based CLOCK algorithm for web-object caching," Master's thesis, EECS Department, University of California, Berkeley, October 2013. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-174.html

[16] M. M. Michael and M. L. Scott, "Simple, fast, and practical non-blocking and blocking concurrent queue algorithms," in *Proceedings of the 15th Annual ACM Symposium on Principles of Distributed Computing (PODC'96)*, May 1996, pp. 267–275.

[17] K. Kim, J. Colmenares, and K.-W. Rim, "Efficient adaptations of the non-blocking buffer for event message communication between real-time threads," in *Proceedings of the 10th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC'07)*, May 2007, pp. 29–40.

[18] M. Yui, J. Miyazaki, S. Uemura, and H. Yamana, "Nb-GCLOCK: A non-blocking buffer management based on the Generalized CLOCK," in *Proceedings of the IEEE 26th International Conference on Data Engineering (ICDE'10)*, March 2010, pp. 745–756.

Design and Deployment of Highly-Scalable Distributed Web Conferencing Systems

Brijesh Kumar[1] and **Constantine Katsinis**[2]
[1] Security Research Group, Rapidsoft Systems, Inc., Princeton, NJ, USA
[2] College of Computing & Informatics, Drexel University, Philadelphia, PA, USA

Abstract - *The current generation of commercially available Web conferencing systems is limited by the number of simultaneous users that can join a Web conference. This number is often limited to a few thousand users at the most. Recently, mconf, a distributed open source-based web conferencing system developed by Brazilian NREN team has been deployed primarily for the educational use. In this paper, we discuss the design of distributed Web conferencing system like mconf that can support a very large number of simultaneous users. Various deployment issues with server load balancing, bandwidth management and streams management are discussed.*

Keywords: Web conferencing, Webinars, Network traffic

1 Introduction

Web conferencing is an interactive online seminar held by a presenter and attended by an audience. Web conferences usually follow an agenda and provide auditory and visual access to the information being shared. Typically, the voices of those attending the Web conferencing are muted during the presentation so that the presenter's message comes across clearly; however, the presenter can take polls or interact with the audience, let people ask questions, and more. Alternatively, the presenter may invite live questions at the end of the Web conference. Similarly, a webcast is a presentation without the interaction.

Web conferences are increasingly being used to deliver marketing material, training and other announcements to captive, geographically dispersed audiences. A number of commercial companies exist that offer commercial Web conferencing capabilities. However, most of these systems only scale to a few thousand users at best. The capacity of GotoWeb conferencing, run by Citrix Corporation, is 1000 users [1]. The capacity of Webex, owned by Cisco Systems is 3000 users (a maximum of 500 participants only can see live, shared video of presenters or panelists, in video-enabled events) [2]. Adobe Connect Server [8] can support up to 1500 users in a meeting, but it is suggested that a Server cluster be used with one server for every 500 concurrent users. Most other commercially available systems do not scale to large number of users.

A typical university may have several thousand students. Similarly, many corporations have tens of thousands of employees. Corporations often need to make policy announcements. To address such a scenario, the companies are forced to use overhead broadcasting systems or invite people to conference rooms or cafeterias. These methods are not desirable since confidential information is often heard by everyone present in the building or there is no interactivity that presenters may like to have. Also, with the cost of education rising, it will be useful to deliver interactive lectures to large number of students across multiple campuses.

Mconf, available from mconf.org, [12] is an open-source multi-conference system built on top of the BigBlueButton open source conferencing system [13]. It is composed of a web portal that provides access to web conferences, shared documents, spaces etc., and a web conference system powered by BigBlueButton which provides the base functionality of the conferencing system. The contribution of mconf team is to add a layer of functionality to load-balance the streaming servers and other load-monitoring communications.

The distributed architecture used in mconf makes it easy to deploy distributed conferencing system using autonomous collection of servers. However, there are still many challenges to deal with if the users are distributed over large geographical areas or in countries with different network bandwidths. We discuss some of these issues later.

In this paper, we discuss a distributed architecture for scalable Web conferencing systems similar to that deployed in mconf. It will allow Web conferencing systems to scale to tens of thousands of users. The rest of paper is organized as follows. In Section 2, we describe the features of a typical Web conferencing system. In Section 3, we discuss the scaling dimensions of web conferencing systems and some benchmarking of single server systems. In Section 4, we present a distributed scalable architecture for deployment of a scalable Web conferencing and conferencing system. In Section 5 we describe bandwidth considerations for scaling Web conferencing systems. In Section 6 we discuss intra-operability and operational issues. Finally, we summarize the paper and discuss future work.

2 Features of a Typical Web Conferencing System

Modern web conferencing systems are expected to provide a wide range of features. Web conferences are just like a conference room-based seminars; however, participants view the presentation through their Web browsers and listen to the audio either through their telephone or via computer speakers. A key feature of a Web conferencing is its interactive elements: the ability to give, receive and discuss information in real time. Most modern Web conferencing systems present a broad range of features. Some features are easily scalable while others are not. Therefore, any architecture that is presented must be able to address most, if not every desirable feature.

All Web conferencing systems provide a similar set of base media broadcast features enhanced by proprietary reporting and event management features. Some of the currently available features are:

- Support for large audiences (100+)
- Registration pages
- Screen and applications sharing
- Document and File sharing
- Chat for asking questions
- Whiteboard
- Collecting attendee feedback
- Polling
- Q&A
- Recordings
- Practice mode
- Voice over the Internet
- Video
- Streaming media
- Reporting (attendance, chat transcript)
- Event management (reminders, planning, invitation)
- Security for User Systems

Therefore, scaling requires the ability to scale the delivery of all the above features.

3 Scaling Dimensions of Web Conferencing

The scaling a web conferencing has two distinct dimensions:

i) Host large number of conferences at the same time: This goal is easily achieved by many commercial systems and is only limited by the number of streaming servers in the network. Each server can only support limited users. So, once that limit is known then scaling is achieved by simply replicating the servers. However, using a single cluster located at one location is not sufficient if the users are distributed over a large area. One of the key challenges of web conferencing systems is to minimize jitter and delays for all receivers to achieve a satisfactory media stream. Adding a suitable load monitoring layer can help minimize the cost.

ii) Increase the limit of users in a single web conference: This is the ability to host a conference and deliver content to a very large number of simultaneous users. This is a much harder problem. It implies enabling a single meeting to be held using more than one streaming server.

Video streaming is a highly demanding application. In a Bluebutton benchmark the following results have been obtained [11] using a server Xeon 3450 quad core 2.66 GHz with hyperthreading:

- 20 voice users and 20 webcams: 30% CPU
- 40 voice users and 20 webcams: 45% CPU
- 60 voice users and 20 webcams: 70% CPU
- 80 voice users and 20 webcams: 90% CPU

It is clear that scaling a web conference for users that come from a large geographical area scan only be done by using distributed server systems.

4 Distributed System Architecture for Scalable Web Conferencing

As mentioned earlier, mconf uses a distributed deployment of servers with a load balancer. In mconf, the web conference servers in the network are shared: every institution that is a member of the network might use any server in it. The server that will be used is decided in the moment a web conference is created and this decision is based in several factors. The server selected might be, for example, the one that is nearest to the user creating the conference, or the server that has more resources available (e.g. CPU and memory) in the moment.

To create a distributed and large-scale conferencing system, one can use a similar architecture. While this architecture works, there are some specific challenges in making a distributed, autonomous systems work. These issues are discussed later.

To scale a single instance of Web conferencing system, as in the case of mconf, we need to add three core components:

i) Replicated web servers with the redundant load balancers: Web servers are used for interaction with the end users in activities such as Web conferencing reservations, finding Web conferencing to join and all other administrative functions like password verification etc. We call these functions Meeting Management (MM) functions.

ii) Replicated media streaming servers: A media server can only serve a limited number of streams. Therefore, replication of media streaming servers is very important for scaling and availability. The key aspect of replicating media servers is to create a chain of media streaming on demand from one server to another.

iii) A protocol to distribute users to various streaming servers and do the dynamic real time load balancing: Since each streaming server has limitations as to the number of streams it can serve, it is important for the system to keep track of the load on each server and update the user clients as to which streaming servers a user client should connect. The users often join and leave a Web conferencing at random times. This means that any load balancing of server and users should be done in real time, requiring the meeting manager managing the Web conferencing to be aware of all users that drop out of a Web conferencing, along with the load on each server so that optimum load balancing can be achieved.

allocation. Two basic steps are followed:

1. Provide a list of possible (and available) media servers nearest to the client's physical location to the user client.

2. Get the measurement of end-to-end delay, and the link bandwidth between client and each media server in this list.

It is not necessary to rely on one particular load balancing method. For example, mconf load balancer implements several algorithms, but only one in use [12]. Some of these algorithms are:

- Select the server with less CPU load;
- Select the server with fewer users;
- Select the server that is geographically nearest to the client.

Since all servers need to implement a similar criterion, it is better to implement a single method. For example, mconf uses the following algorithm [12]: Order the servers by proximity to the client. If there is more than one server in a ~300km radius, select the one that has less CPU load in the moment. It does not consider servers that are not up or that are not responding properly.

After considering the data supplied by the client, a media server is assigned to the client for joining a live stream. Figure 1 shows the architecture and its various components.

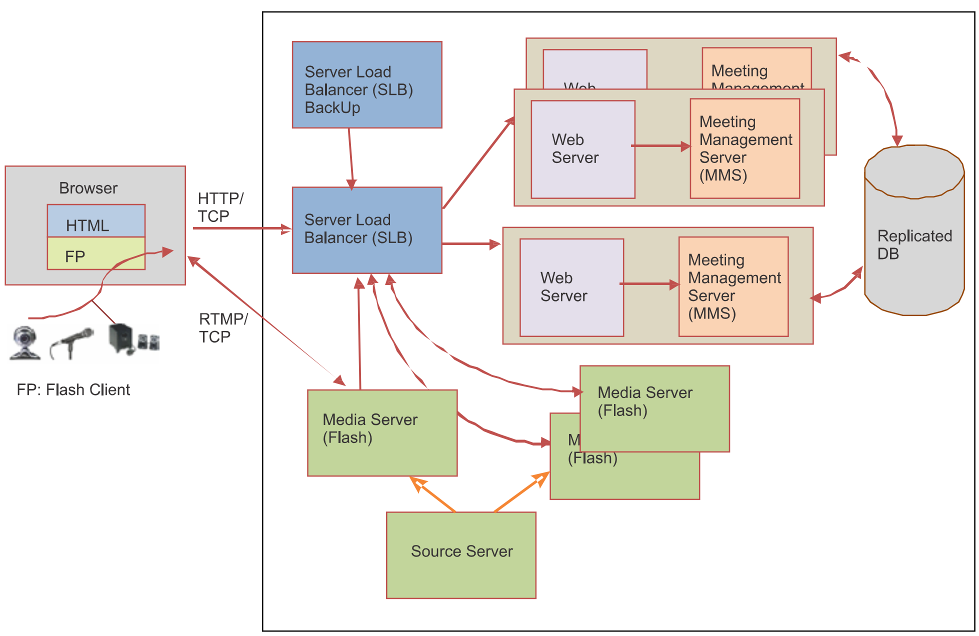


Figure 1 Architecture and Design of Scalable Conferencing System

The load balancing process interacts with the client to collect the data that it considers in media server

In this architecture, clients (web clients and Web conferencing clients) connect to a farm of Meeting

Management System (MMS) servers via a server load balancer (SLB). The SLB is used to distribute incoming HTTP requests amongst a pool of MMS Servers. The SLB can be configured to allocate requests in a round robin fashion or based on MMS server load.

The SLB monitors the MMS servers, and removes them from the pool in case of failure. The SLB itself has a backup in case of failure. The use of the SLB thus promotes availability, but it also promotes scalability, by making it easy to add new servers without impacting any system components beyond the SLB. In addition, all MMS servers are stateless, meaning that any server can take over from any other server in case of a failure.

The Media Servers implement the communication for individual meetings: screen sharing, VoIP, video, file and chat data. The MMS Server makes the decision as to what Media server is to be used for a given meeting. MMS Servers load-balance Media servers, and remove Media servers from their available pool should Media servers become unavailable. This provides both scalability in the number of meetings that can be hosted, and failover capability should a Media server cease to operate correctly.

Finally, conference data (such as account information, schedule information, meeting status and billing data) is stored on a database server. To provide failover capability, a backup database server mirrors the database server in real time.

4.1 Architectural Tradeoffs in Distributed Media Servers

Any architecture embodies tradeoffs. Distributing servers at many locations may not be possible because of cost considerations or lack of availability of high speed bandwidth. In Web conferencing systems, it is important to keep the cost down. Therefore, if it is possible to use servers that can be provisioned on demand then the cost can be reduced without affecting scalability. However, centralizing all servers at one physical location may lead to degrading of the media experience for some users. Therefore, a server allocation must consider the delay between the clients and the servers in considerations when assigning users to a particular media streaming server.

4.2 Integrating Telephony Voice With Media Servers

While most users join a web conference using a computer system with attached microphone and speakers, some users may like to join a phone-only connection. Sometimes, even users with PCs and desktops may like to have their audio delivered via a phone connection.

The integration of VOIP conferencing services and the Media server is done using RTMP. VOIP servers such as open-source Asterisk or Free Switch are integrated with media streaming servers. Their phone switch provides a RTMP channel to the media servers. The quality of voice using a phone will depend on the general call quality between the VoIP switch and the phone.

4.3 Dealing with Difference in Bandwidths in Different Areas

One of the challenge in delivering a web conference to a large number of users is to deal with different bandwidth available to users. For instance, the users in the US have access to 25 Mps connections at home using fiber or cable. However, users in developing countries only have access to a 100 to 500 kbps links in their homes. This means that a video stream that is easily received and played by the users in the US cannot be received by users in other countries due to lack of bandwidth.

This problem can be addressed to some extent by the use of scalable video encoding. Scalable Video Coding (SVC) is the name for the extension of the H.264/MPEG-4 AVC video compression standard. SVC standardizes the encoding of a high-quality video bitstream that also contains one or more subset bitstreams. A subset video bitstream is derived by dropping packets from the larger video to reduce the bandwidth required for the subset bitstream. The subset bitstream can represent a lower spatial resolution (smaller screen), lower temporal resolution (lower frame rate), or lower quality video signal.

This allows users in the low bandwidth areas to display only one layer (poor quality) and the users in high bandwidth area to display all the layers (high quality video).

5 Media Server Bandwidth

A media streaming server is the most important part of the Web conferencing system. The scalability of a media server can be computed by estimating the bandwidth required for Web conferencing.

We describe a method for estimating bandwidth for an application such as video conferencing or Web conferencing as described in [6]. A simple formula for calculating bandwidth is used for both the server bandwidth (which is required to determine the approximate number of server instances that the system needs) and the client bandwidth (which should be calculated as part of overall bandwidth strategy to insure good quality of service to each client).

Clients connect directly to a Web conferencing Media Server which handles on demand and real time video, audio, and data delivery. As the number of clients increases, the resulting high bandwidth requirements may potentially increase the latency during the Web conferencing. As discussed earlier, Web conferencing Media Servers need to be provisioned in the proper number and at the proper locations to ensure that the bandwidth requirements are satisfied.

In the case of Web conferencing, the key parameter is the maximum number of clients attending at each major location. Since clients usually need to register before the Web conferencing, the organizer can easily estimate attendance. The following analysis examines the bandwidth requirements at a single location relying on experimental results from [9] and [10].

Bandwidth requirements depend on the type of usage necessary for the Web conferencing such screen sharing, audio, webcam video and streaming video. Experimental results in [9] indicate that typical webcam video bandwidth per client is 239 Kbps and 501 Kbps depending on whether the client is using a DSL or LAN connection respectively. It is assumed that the Web conferencing Media Server is aware of the type of connection the client uses and adjusts the webcam video parameters for each client accordingly. Specifically, the DSL bandwidth results from webcam video of 640x480 at 15 frames per second (fps) and compression quality of 85; the LAN bandwidth results from webcam video of 640x480 at 20 fps and compression quality of 90.

Similarly, experimental results in [10] indicate higher typical webcam video bandwidth per client of about 900 Kbps for LAN connections due to higher-quality compression despite a lower resolution of 640x360. Results in [10] also show that webcam video bandwidth is highly dependent on whether the subject is moving. The bandwidth of 900 Kbps was observed with a slowly moving subject while a stationary subject resulted in a significantly smaller bandwidth. ([10] reports a reduction of bandwidth from 308 Kbps to 58 Kbps using a lower video resolution of 320x180).

Depending on quality and frame rate, the bandwidth needed for screen sharing has been observed to be between 138 and 164 Kbps for a LAN connection and slightly less for a DSL connection [9]. VoIP audio bandwidth was observed at 44 Kbps for a LAN connection and 22 Kbps for a DSL connection.

Bandwidth for streaming video plus audio was observed to be approximately 1200 Kbps (with small variation depending on the file format).

The total bandwidth requirements per client are approximately 700 Kbps for a LAN and 400 Kbps for a DSL connection, without the use of streaming video. Given the experimental results in [10], a more realistic estimate of the needed bandwidth for a LAN connection would be 1000 Kbps. Assuming that a Web conferencing Media Server has a 1 Gbps network connection available, the total number of clients that could be supported by a single server would be approximately 1000 clients, all using LAN connections, to 2000 clients, all using DSL connections. If streaming video (at 1.2 Mbps) is also used by the Web conferencing, then these bandwidths are reduced to approximately 450 clients, all using LAN connections, and 600 clients, all using DSL connections. For a mixture of LAN and DSL clients, the required bandwidth would be between these two endpoints.

Even if the network connection bandwidth available to a Web conferencing Media Server at a single location Web conferencing Media Server is 1 Gbps, conservative estimates suggest that a single server may be able to support up to 500 clients. When network connection bandwidth allows more clients to be connected, then two or more servers would be utilized to support all the connected clients.

6 Intra-operability and Operation Issues

Many conferencing systems are available today but there is no interoperability between them. In order to accomplish interoperability among different vendors, common standards must be followed. The two most prominent standards are H.323 (ITU-T H.323, 2009) and SIP (Session Initiation Protocol), and both are used worldwide to allow interoperability among different systems.

The basic idea of these protocols is to have some common messages to initiate the session and exchange capabilities. Using the information exchanged, the different parties in the conference can agree on the same group of audio and video codecs, as well as the data protocol, allowing the interoperability. In other words, if a distributed system on geographical large scale is to be deployed all systems must use the same software for signaling and load balancing.

This leads to another problem common to all distributed systems. Since systems interoperate and connect with one another, the software change in one server affects all other. If the network consists of 1000 nodes, changing software on all 1000 nodes at the same time requires special software tools and logistics.

7 Conclusions

In this paper we presented various aspects of a Web conferencing and conferencing system. We also proposed a system architecture that can be used to deploy a highly scalable Web conferencing delivery architecture. Web conferencing delivered to large, geographically dispersed audiences, can benefit by relying on a distributed architecture of media servers provisioned on demand and conveniently located so as to reduce cost and maintain scalability. Our analysis shows that individual media servers can support between 450 and 2000 clients depending on video requirements of the Web conferencing and the network bandwidth available at each client. Therefore, a small number of media servers provisioned at the proper locations can readily support large audiences.

The system described can be built using a Flash Media Server (Adobe Flash Server, Red5, Wowja Flash Media Server), and adding client software and load balancing modules. A simple prototype can be developed using mostly open source components and integrating open source VoIP telephone systems such as Asterisk or FreeSwitch.

8 References

[1] Goto Web conferencing, Citrix FAQ, http://www.citrixonline.com/collaboration/images/pdf/GoToWeb conferencing_QA.pdf

[2] Webex Web conferencing, Webex Count, http://www.webex.com/products/Webconferencings and online events.html

[3] Bin Yu, Yong Rui , AWeb Conferencing Systems for Ubi Media Collaboration: Achievements and Challenges, First IEEE International Conference on Ubi Media Computing, July 31 Aug. 1 2008 , Page(s): 279 283

[4] Kundan Singh, Carol Davids, AFlash based Audio and Video Communication in the Cloud, Research Note, (2011), http://arxiv.org/ftp/arxiv/papers/1107/1107.0011.pdf

[5] Chris Hock, ACalculating your Bandwidth and Software License Needs for Flash Media Server 2, Adobe Research Note January 19, 2007

[6] J.G. Apostolopoulos, W. Tan, and S.J. Wee, AVideo Streaming: Concepts, Algorithms, and Systems, HPL-2002 260.

[7] A Real time messaging protocol, Whitepaper, Adobe Systems, Jun 2009, http://www.adobe.com/devnet/rtmp.html

[8] Adobe Connect Server, http://www.adobe.com/products/adobeconnect.html

[9] Adobe Connect Technical Overview, http://www.getconnect.com/wpcontent/uploads/Adobe Connect 9 Technical Guide.pdf

[10] ACisco WebEx Network Bandwidth, White Paper, June 2013, http://www.cisco.com/c/en/us/products/collateral/conferencing/webex meeting center/white_paper_c11 691351.pdf

[11] http://www.inf.ufrgs.br/prav/mconf.htm, NREN performance Presentation.

[12] mconf, http://mconf.org project home site and project documentation.

[13] BigBlueButton, http://bigbluebutton.org/ Project Documentation.

Disambiguating Cloud Security

T. Ray Campbell

Computer Information Systems Department, Livingstone College, Salisbury, NC, USA

Abstract - *This work defines a vision paper that captures the intrusion parameters for a cloud network operational within the author's own University environment. As a contribution a threat and security model that mitigates these concerns are provided within the context of a case scenario.*

The delivery of data services is becoming more pervasive and as such there is concurrent need for ensuring efficiencies in delivery and in the security of the process. Security for physical network equipment is no longer as worrisome for administrators; it's the protection of the data and exposure to sensitive information that is the most important aspect for data security. Cloud computing is evolving to solve the equipment issue but the data security is yet to be mastered. Motivation for this work is the author's design of a threat model that show constructing formal arguments on how these threats can be mitigated as the specific paper contribution.

Keywords: Cloud, security, security threats, threat model

1 Introduction

Cloud computing is defined in the National Institute of Standards and Technology (NIST) document as "an ubiquitous, convenient and on-demand computing environment that allows for shared, on-demand, and easily configurable resources and services that can be readily provisioned and released with minimal provider interaction or management effort. The cloud-computing model promotes availability and is delivered as a software, infrastructure, and platform as a service model"[1].

No longer do network administrators; data owners in organizations and network security infrastructure engineers have to worry about physical equipment in clouding infrastructure [2]. The important consideration is the protection of the data. Critical to this is the data lifecycle and the probability of exposure of sensitive information.

The cloud in its ubiquitous delivery mode is radically changing how information and services are delivered [3]. This ubiquity of cloud operations presents various challenges for security administrators while opening a haven for nefarious activities by cyber-criminals.

There are naturally multiple security concerns about data protection but there are also concerns about the actual vulnerabilities and protection of the virtual machines that make up the cloud infrastructure. Penetration testing data have shown multiple attack vulnerabilities existing in public clouds ranging from obsolete versions of Web services software, exploits of HTTP protocols, and virtual machines (VM) images [3].

What does 'disambiguating cloud security' mean in security for cloud computing? In the context of this presentation, disambiguating means the deeper relationship between that which is presented on the surface and which actually exists. Smith and Martins [4] speak of a level where "dependency" that is integral for learning algorithms and parsing are used to extract information for predictive behavior. These activities are key components for data forensics, which are usually an after-event activity. This principle of disambiguating cloud security is expected to allow for a predictive and preemptive mode of continuous monitoring and analysis to mitigate and correct intrusive behaviors in the cloud.

1.1 Cloud Architecture

Cloud computing as presented by the National Institute of Standards and Technology (NIST) [1] is an ubiquitous, convenient and on-demand computing environment that allows for shared and easily configurable resources and services that are readily provisioned and measured based on needs. The characteristics inherent in cloud computing provide for economies of scale, flexibility, increasing productivity in infrastructure services, a greater ability to focus on core competencies and sustainability, and transparency in delivery and control of the services that are being provided [5].

Table 1: Characteristics of Cloud Computing

Characteristics	Description
On-demand service	IT services are readily available on demand with increased flexibility and availability
Network Access	IT services are available via networks independent of the user end device via high performance broadband technologies
Resource sharing	Resources are provided to multiple consumers via technologies of virtualization and multi-tenancy; scalability and adaptation to changing needs
Rapid Provisioning	Elasticity in provisioning resources rapidly and releasing them without manual intervention, an added benefit for disaster recovery and load balancing operations through mirroring; efficiency
Measured Services	Services consumed are easily measurable and billed based on consumption, a 'pay as you go' / 'pay per use' methodology.

Source: Based on NIST Definition of Cloud Computing – Publication 800-145 [1]

The operation models for cloud services delivery are distinguished as Infrastructure as a Service (IaaS), Platform as a service (PaaS), and Software as a Service (SaaS). These models provide individual benefits based on the services required by the consumer and the IT architecture that the services reside on. Delivery of the services follows a deployment model of a private cloud, a public cloud, a hybrid cloud, or a community cloud as depicted in Figure 1.

Figure 1 - Visual Reference Model of Cloud Deployment

Layer	Cloud Computing Components
Five Characteristics	On-demand self-service; Broad network access; Resource pooling; Rapid elasticity; Measured Service
Three Delivery models	IaaS; PaaS; SaaS
Four Deployment models	Public; Private; Community; Hybrid

2 Problem Formulation

The rapid expansion of computer networks mainly due to the Internet has created security issues related to data confidentiality, integrity, and availability. These issues have manifested in numerous attacks and threats that have increased significantly with frequency and variety. The nature of these attacks have resulted in intrusive behaviors that have caused denial of services to users that force the performance of networks to carry out the tasks that they were intended, thus affecting the overall performance of networks.

2.1 Threat Model

Security in the cloud forces the user to trust that which he has no control over and limits that which he can do, even though from the provider's perspective, there are benefits to be derived from this paradigm. Dhage, et al [6] relate this to the a transportation system that forces the consumer to be dependent on availability of transportation even though he has no control over the scheduling and thus little or no alternative other than to trust that the scheduling will adequately meet his needs. It is a touch and go situation.

2.2 The Model

Security threats can take many forms. The assumption presented here is that the attacker will either be an insider or and outsider, and that the main goal is to exploit any weaknesses or vulnerabilities that may exist in the virtual networking system [7]. Varying threat models have been promulgated as cited in [8], but the emphasis here will be focused on data and virtualization threats.

2.2.1 Data Threats

Concerns over privacy and security of information systems and more so securing data in the cloud is an ongoing problem that has prevented many businesses from fully accepting cloud platforms. On-going research to develop and implement secured solutions has become of paramount importance for business sustainability, upholding user privacy, preserving data integrity, and instilling trust.

The integrity of data can be easily compromised and is worsened with theft or loss. This problem does exist in a physical environment but can become more acute in the cloud environment. This is due in part to the extensive amount of data that a service provider is managing; as the amount of data increases, the propensity for exploits is likely to increase significantly. The fact that communication in the cloud is primarily via an Internet connection, there will always be the probability for data leakages [8], and this therefore means that enhanced methods of authentication and encryption should be built in in order to minimize the prospects of leaked data being accessible to the perpetrators who may be outside attacks or insider threats. Attacks on data systems are inviting to intruders and as more personal and business financial and other proprietary data are stored online, the vulnerabilities are more prevalent.

Threats to the data in the cloud may comprise of run-time execution, malicious threats such as denial of service / distributed denial of service (DoS/DDoS), natural system threats such as accidents and the adversarial types that include insiders. Natural system type threats like accidents can be predicted through estimation and probability theories, and with redundancies put in place for failover, an acceptable threshold can be developed. However in the case of the adversarial threats and attacks such as DDoS/DoS, it becomes highly unlikely that a predictably point of acceptance can be determined. The client organization (who uses a cloud provider) has no knowledge of, or interference with the personnel who interacts with or manages their data. Malicious attacks on the data can result in devastating impacts. What, therefore are the solutions? These rests in part on trust; trust that the provider will apply appropriate legal and operational measures to assure that vulnerabilities are minimized and when present, they are easily and quickly contained [9].

There are other broad categories of threats to data that impact the network along with the cloud environment. Bishop [10] lists four such threats that affect information systems:

disclosure, deception, disruption, and usurpation. Following along the same vein, Amoroso [11] identified three categories that impact security, namely disclosure, integrity, and denial-of-service. Each threat impacts the information system differently and as such demand a different approach to mitigate.

A *disruption threat* in this context refers to any event that prevents or impairs the normal use and function of a computer system. This could be physical, meaning any action or threat resulting from the attack to the physical objects required for operation of the system, such as power failure, dislodging or cutting of cables, whether intentional or unintentional, fire or flooding in a computer room. To help mitigate against these threats organizations must restrict physical access to the system, to cabling and networking closets, electrical supply, use cameras, guards and locked doors.

A *disclosure threat* involves unauthorized access to information, which can take the form of either technical or non-technical, and targeted at personal information for identity theft, as well as extortion.

A *deception threat* takes the form of gaining access to a system in order to disclose, alter or disrupt services by impersonating a legitimate user. This misrepresentation could be social engineering, spoofing or falsifying messages or data packets.

Integrity threats are often represented through modification or altering of data to obtain a different result than that which would normally be if unaltered. Actions of deleting records in a database, altering academic grades, deleting criminal records, altering payroll records to reflect significantly different amounts that could affect tax computation, are only some of the impacts that present in integrity threats.

Usurpation threats refer to the acquisition of privileges that should not be available to an unauthorized user. These acts of threats against the security of the data are all intended to violate all aspects of access control mechanisms that an organization may have in place, and with all measures that may be instituted, there is still the element of trust. One must be able to trust those that are given authority to access a data system.

Trust models in the cloud are arguably absent and this is undeniably an issue that will grow as cloud architecture and scalability grows. As interactions with cloud providers, providers' reputations, user knowledge, and experiences grow; trust elements will conceivably grow [12].

Another security threat that exists within the cloud for data surrounds the inevitability of the data owner moving from one cloud provider to another. What guarantees exist that all data will be transferred to a new provider, and what assurances are there that any remnants will be erased from the initial provider's servers, thus eliminating any adverse effect after the fact?

2.2.2 Virtualization Threats

The cloud-computing paradigm provides high scalability and flexibility with service on demand. As such it is highly dependent on a dynamically virtualized environment that uses hypervisors in order to satisfy the infrastructure; Infrastructure as a Service (IaaS).

The hypervisors are running from physical servers and these servers are part of the network. These threats are all part of network security threats that will impact the infrastructure as they do to the data. Therefore threats such as DoS, DDoS, IP spoofing, ARP spoofing, and multiple other threats that affect physical servers are also threats to the hypervisors [6].

Virtualization does not ensure isolation; the cloud architecture offers multi-tenants on its infrastructure, so it is evident that a compromise of the hypervisor will likely affect multiple tenants. In addition, the hypervisors can allow attackers the privilege of gaining access and control of the shared platform as indicated in [8]. The fact that the hypervisor acts as the emulator of the underlying server hardware and provides an interacting environment for the virtual machine (VM), a malicious VM can be used to attack the hypervisor or any other VM that may be running from that infrastructure. This promulgation is reinforced in [13].

The vulnerabilities of the hypervisor is an attackers paradise; the main operational reason for compromising the hypervisor is to gain access and control of the underlying layers that are needed to provide functionality to the VMs installed on the physical server [14]. An attacker gaining access and compromising the hypervisor could effectively use the VMs on that infrastructure to launch a DDoS attack on neighboring servers and VMs.

The VMs themselves are also vulnerable. Since the architecture of the cloud infrastructure contain many and varied VMs all with different images and applications, and because the cloud is effectively multi-tenancy, an attack or compromise of one VM whether it's active online or not, could effectively compromise the others, on the same physical server.

The extent of the security threat [14] is to a large degree based on, and due to the complexity of the infrastructure and the dynamics of the use of the cloud. The cloud IaaS layer facilitates multiple users from various geographical locations who all participate in a sharing domain that could foster data leakages that are at risk for VM attacks [15]. At the same time users have very little control over, or awareness of the location of the physical server(s) that their data is stored on. The cloud service providers themselves are not aware of the contents of the VMs or the applications that they are hosting. This therefore is an opening for multiple security threats.

There are other threats that arise as a direct result of the virtualization for the IaaS and they include:

- Complexity in the workload: the voluminous nature of network traffic that runs on the cloud servers are increasing as the adoption rate increases. This therefore could result in complexity in managing the workload of each server / cloud provider.
- Loss of Control: users are not aware of the location of their data and service neither the whereabouts of the physical servers. Likewise, the service providers are not aware of the contents on the VMs.
- Network topology: the architecture of the cloud is very dynamic. Due to this, providers will have to continually create and remove VMs in order to accommodate the workload and traffic. When VMs migrate from one server to another without a pre-defined network topology, there is a likelihood of creating security vulnerability.
- Single Point of Access: the cloud functions in a virtualized environment and the virtualized servers have limited physical access via network cards (NICs) to the VMs. Consequently; this is a critical point and vulnerability where a compromise could be launched on the hypervisor.

The threats enunciated above are only some of vulnerabilities that pose serious disambiguation in understanding security in cloud computing, which all pose challenges for protecting confidentiality and integrity. The main purpose and goal of an attack in the cloud, is to disrupt or deny full functionality and availability of cloud services.
This in effect would be a DoS attack on the environment of which significant impact would be felt. The extent of this is directly proportional on the business operations and criticality of the organizations that use cloud services. To mitigate against these eventualities, strict access methodologies and monitoring must be implemented.

2.3 Proposed Countermeasures

Threats and attacks to the computing infrastructure such as DoS and DDoS are designed to exhaust and tie up network resources to the point of crippling productivity. This eventuality can however be mitigated if proactive and detective measures are in place. In addition, safeguarding data in the cloud is important, and some of the measures are:

- Encryption: data stored in the cloud should be encrypted. The trust level of the implemented cloud security [16] may require limiting the risk of storing data in an unencrypted form.

- Trust: use a dynamic approach to granting access to users based on their organization's role. Some parameters to consider may be location based, devices, and resources needed and requested and configured within existing policies framework.

- Install intrusion detection systems [14] to perform pre-emptive detection through detective monitoring of all incoming and outgoing traffic for abnormal activities and analysis, prevention, and effective incident reporting.

- Determine sensitivity of data in line with organizational roles based and trust levels.

- Protect the virtualized infrastructure by applying operating system (OS) hardening principles, establishing trust zones, and isolate VMs so that they don't interfere or conflict. Provide isolated unprotected VMs to display 'bot net' decoy behavior and strategies.

- Perform risk assessments of applications to be deployed in the cloud before making them available for user access.

- Evaluate and apply network architecture for configurations that assure withstanding against attacks, apply defense-in-depth approaches and mechanisms, evaluate authentication and access control methodologies and policies and eliminate any overlaps that could render the system vulnerable.

- Install active monitoring sensors to monitor performance that work in concert with IDS to relay information for detection of zero-day threats [15].

3 Future Works

There are many other areas of security vulnerabilities that confront the cloud arena in particular, insider attacks that are very important to assuring cloud security. Driven by the limitations of existing security measures as a paradigm for the cloud, and as a continuation to this position paper, the author intends to pursue development of a security model for cloud computing. This model would present the need for a Security as a Service (SECaaS), and would be promulgated by the requirements and principles that underlie the infrastructure as a service, and the need for a secured environment.

4 Conclusions

This paper sought to introduce a discussion on the disambiguation surrounding cloud security. Cloud security is an important topic that creates ongoing dialogue and research to find the best combination for security; confidentiality, integrity, availability, access, and trust, with proper defense strategies. As cloud computing systems continue to grow perversely, secure preservation and safe access to information through hardware and software solutions such as detection and prevention systems, encryption, and strong authentication mechanisms are needed [17]. Securing the systems and data are ongoing activities that require tools to perform early detection, identification, and mitigation of malicious activities.

5 References

[1] P. Mell, & T. Grance, (2011). The NIST definition of Cloud Computing: National Institute of Standards and Technology (NIST) Special Publication 800-145. Retrieved at http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf December 7, 2013.

[2] D. Mitropoulos, (2013). Data security in cloud environment. *XRDS, Spring 2013*, 19(3).

[3] M. Simmons, & H. Chi, (2012). Designing and implementing cloud-based digital forensics hands-on-labs. In *Proceedings of the Information Security Curriculum Development Conference 2012*, Kennesaw, GA, Oct.12-13, 2012.

[4] N. Smith, & A. Martins, (2013). Linguistic structure prediction with the sparseptron. *XRDS Spring 2013*, 19(3)

[5] M. Kaiserswerth, (2013). White Paper on Cloud Computing, Swiss Academy of Engineering Sciences, 2013. Retrieved at: http://www.satw.ch/organisation/tpf/tpf_ict/box_feeder/2013-02-06_SATW_White_Paper_Cloud_Computing_EN.pdf December 2013

[6] S. Dhage, B. Meshram, R. Rawat, S. Padawe, M. Paingaokar, & A. Misra, (2011). Intrusion detection system in cloud computing environment. In *Proceedings of International Conference and workshop on Engineering Trends in Technology (ICWET)* 2011, Mumbai, India

[7] C. Chung, P. Khatkar, T. Xing, J. Lee, & D. Huang, (2013). NICE: Network intrusion detection and countermeasure selection in virtual network systems. *IEEE Transactions on Dependable and Secure Computing*, 10 (4) July/August 2013

[8] K. Lourida, A. Mouhtaropoulos, & A. Vakaloudis, (2013). Assessing database and network threats in traditional and cloud computing. *International Journal of Cyber-Security and Digital Forensics (IJCSDF)* 2 (3), 2013.

[9] M. Burmester, (2013). Trusted clouds. In *Pre-Proceedings of International Workshop on Trustworthiness, Accountability, and Forensics in the Cloud* (TAFC), Malaga, Spain, June 6-7, 2013 ISSN: 2078-2247

[10] M. Bishop, (2002). *Computer Security: Art and Science*. Boston, MA: Addison-Wesley.

[11] E. Amoroso, (1994). *Fundamentals of Computer Security Technology*. Englewood Cliffs, NJ, Prentice Hall

[12] S. Thorpe, (2010). Modeling a trust cloud context. In *Proceedings of ACM PKIM'10*, Toronto, Canada, October 30, 2010

[13] J. Szefer, E. Keller, R. Lee, & J. Rexford, (2011). Eliminating the hypervisor attack surface for a more secure cloud. In *Proceeding of ACM CCS'11*, Chicago, Illinois October 17-21, 2011

[14] B. Mansukhani, & T. Zia, (2012). The security challenge and countermeasures of virtual cloud. In *Proceedings of the 10th Australian Security Management Conference*, Perth, Western Australia, December 3-5, 2012

[15] A. Kumar, C. Srinivasulu, B. Sudeep Kumar, & D. Priyanka, (2013). Emphasis and emerging trends on virtualization of cloud infrastructure with security challenges. *International Journal of Computer Trends and technology (IJCTT)*, ISSN: 2231-2803, 4(8), August 2013.

[16] M. Lemoudden, N. Bouazza, B. El Ouahidi, & D. Bourget, (2013). A survey of cloud computing security overview of attack vectors and defense mechanisms. *Journal of Theoretical and applied Technology, ISSN: 1992-8645*, 54(2), August 2013.

[17] A. Patel, M. Taghavi, K. Bakhtiyari, & J. Júnior, (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of Network and Computer Applications 36* (2013) 25-41

Characteristics of Online Social Services Sharing Long Duration Content

Hyewon Lim[1], Taewhi Lee[2], Namyoon Kim[1], Sang-goo Lee[1], and Hyoung-Joo Kim[1]
[1]Computer Science and Engineering, Seoul National University, Seoul, Republic of Korea
[2]BigData Software Platform Research Department, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea

Abstract – *Recent advances in web technologies have brought a boom in online social services. Online social services encompass a broad spectrum of how to cater to varying web users' interests, social content sharing being one of them. Studies show that users of social content sharing services tend to focus on the items themselves, rather than communicating with others at all. Unlike social networking services, which generally exhibit a healthy degree of user interaction, user interaction within social content sharing services shows much variation, from nonexistent to very active. Consumption time of target content has also shown to be an important factor in determining the various features of content sharing services. In this paper, we analyze social cataloging services and compare them to the YouTube network in* [1], *to figure out the characteristics of networks which contain content that requires a considerable amount of time to fully digest, such as films or books. We find that our dataset shows a higher level of homophily, reciprocity, and a lower level of assortativity than those of the YouTube network. Furthermore, we study features that affect users' selection of* new items. *Our results show that interest similarity matter, the genre of books in particular.*

Keywords: social content sharing service, content consumption time, social cataloging service, influential features

1 Introduction

Over the past several years, online social services have grown at an unprecedented rate, and have served as a catalyst in the explosion of online media content. Unlike the past when only a small number of people were capable of creating media for the public to consume, we are already in an era of a user created content deluge. Online social services have played an important role in allowing creative users to share their content and find an audience. By supporting user activities, they contribute to enhancing interconnectivity, self-expression and information sharing [2]. Online social services can be classified by their differing aims, system components, and the behavioral patterns of users. Among them, social content sharing services deal with target content such as videos for YouTube, books for LibraryThing, and images for Flickr. The main purpose of social content service users is to share their items and to find other items that match their interests, rather than to make a relationship or communicate with others.

Social content sharing services look very similar to each other, but each of them has different characteristics depending on its target content. Music, video clips, or photos have short consumption times and users can usually access them within the service; however, movies or books, which we will refer to as long content, require relatively longer consumption times, spanning from a few hours to several days, and services dealing with long content usually provide users with only the meta-data and perhaps a fraction of actual content, such as a sample teaser or chapter. Because of the fundamental difference in how short and long content are provided in social content sharing services, content consumption time can affect user's content selection and communication. For long content, users will select items carefully based on personal preference; they also tend to express their opinions through reviews or scores rather than casually conversing with others. That is, long content users maintain their relationships by reading news feeds about their friends, without any interactions.

Although communication in social content sharing services is less active compared to that in more general social services, a user's item selection is not independent of others' behaviors and must be influenced by exposure to social opinion and sentiment, even by a miniscule amount, because of the various social features typical of such services. Social Influence is defined as change in an individual's thoughts, feelings, attitudes, or behaviors that results from interaction with another individual or a group [3]. In an online social service, users influence is based on the trust in users or contents, either through relationships or sharing items. Understanding the features that propagates influence is important in analyzing, recommending, or advertising items and users.

In this paper, we analyze long content services LibraryThing, a social book cataloging service, and Userstory Book, a similar service primarily used in Korea. As the names suggest, the target items are books. We compare said services' data to a YouTube analysis [1] in regard to assortative linking, reciprocity, and homophily, with varying content consumption times. We find that the assortativity of social content sharing services with long content is low, but the level of reciprocity and homophily is high. We also analyze social features such as interest similarity and user behavior in order to figure out which ones affect users' item selection. The most influential feature is the interest similarity among users; in our dataset the genre of a book plays an important role in making relationships and selecting new items. We believe that our work is the first study to analyze social content sharing services in terms of varying content consumption times.

The remainder of this paper is structured as follows: We briefly introduce some related work in section 2. Our data set and experiment results to compare with [1] are presented in Sections 3 and 4. Section 5 shows which features play an important role in users' item selection, followed by the conclusion and future work in Section 6.

2 Related Work

As online social services become more popular, researchers try to analyze them to understand their key characteristics. Some researches try to categorize the services; its result differs based on various perspectives. Kaplan *et al*. [4] sort social media services into six categories by intimacy and immediacy. In this research, social content sharing services are classified as "Content Community", with medium social presence and low self-presentation. In [5], online social services are categorized into four groups by formality and interaction. They put social content sharing services in the "Cooperation" group, which has low formality and high interaction.

Much effort has been made for analyzing the static and dynamic features of social content sharing such as network structure and user behavior. User behavior is an especially important feature in understanding phenomena present in social networks, such as social influence and user similarity. Mislove *et al*. [6] analyze the structural characteristics of four popular sites: LiveJournal, YouTube, Flickr, and Orkut. Wattenhofer *et al*. [1] study the social network in YouTube. They show the differences between the YouTube network and traditional online social networks using three features: assortative linking, reciprocity, and user homophily. They also show the dichotomy of 'social' and 'content' activities and examine said activities' popularity in YouTube. The analysis of social cataloging services performed by [7], [8] both use an aNobii network. In [7], they investigate structural and evolutionary features and mine geographical information. Tang *et al*. [8] study the reading diversity of users using five similarity measures. Calculation of the interest similarity of users by using relationships is performed in [9]. They show that the similarity decreases with the weakening of connection strength between users. Crandall *et al*. [10] study the role of user interactions between similarity and social influence. They find that social interaction is both a cause and effect of similarity and social influence. According to their research, users show a sharp increase in similarity immediately before their first interaction; after the interaction, their similarity is increases slowly.

There has also been much effort to model social network phenomena. In [5]–[8], the researchers try to model social influence using user behavior. Yeung et al. [11] propose a probabilistic model for user adoption behavior to capture implicit influence in social content sharing services. In [12], [13], they research topic-based social influence.

3 Dataset

Social cataloging services are suitable for figuring out the characteristics of long content services because it naturally takes longer to read a book than to view a photo or watch a short video clip.

Users make their own reading lists, rate the books, and write reviews. Naturally, they also make relationships, comment on other's pages, and join communities much like when using online social network services. In this paper, we use LibraryThing and Userstory Book for the dataset. Table 1 shows the summary of our dataset.

Table 1. Data Summary

	Library Thing	**Userstory Book**
The number of users	108,221	12,933
The number of relationships (unilateral)	302,728	13,591
The number of relationships (reciprocal)	225,783	7,582
The number of books	13,285,867	100,168
The number of comments	161,340	2,181

3.1 LibraryThing

LibraryThing is one of the most famous social cataloging services launched in 2005. It has almost 1.8 million users and over 80 million book information so far. The user relationship is unilateral; they do not need to get consent to be connected. In addition to the functions described above, it is possible to tag the book in the list of their own.

In a social cataloging service, user behavior related to books is more important than user relationships; therefore, it is difficult to crawl data using some users as a seed like in social network services. We choose users who have one of three books as a seed, the books being "The Casual Vacancy" and "The Hunger Games", which are the most popular books in our crawling period and "Wuthering Heights", which has been loved for a long time. The reason we select "Wuthering Heights" is to avoid collecting users in a limited age group. We collected the data from January 23rd, 2013 to January 30th, 2013 using breadth first search. The dataset contains 108,221 users, 302,728 unilateral relationships, 13,285,867 book entries, and 161,340 comments.

3.2 Userstory Book

Userstory Book is a social cataloging service in Korea launched in 2009. It has almost 20,000 users and over 180,000 book entries so far. The relationship is unilateral like that of LibraryThing. The users cannot tag their books on their own lists, but they can sort the status of the books into three groups: "plans to read", "currently reading", and "already read." We collected the entire data until May 8th, 2012. The dataset contains 12,933 users, 13,591 unilateral relationships, 100,168 books, and 2,181 comments. The size of the dataset is remarkably small compared to the LibraryThing dataset; however, it is important to understand what the characteristics of the entire network of the social cataloging service imply. We will also show that the Userstory Book is suitable for analyzing because it has a similar structural tendency to the LibraryThing dataset despite the smaller size.

3.3 Analysis of network structure

In both services, user interactions occur if a user leaves messages on others' pages (or wall in Userstory Book) or replies to reviews. In fact, most users rarely comment on others' reviews; hence, we make an interaction graph based on the comments on personal pages. If there is an edge from user u to user v, it means u leaves a message on v's page. We ignore the replies of the messages and the messages written to oneself. The interaction graph of LibraryThing consists of 29,989 users and 94,209 interaction edges, and 11,524 users have one or more reciprocal edges. In the case of Userstory Book, the interaction graph consists of 827 users and 1,287 interaction edges and 321 users have one or more reciprocal edges. It shows that 23% and 6.7% of the total users have interactions with others. Of these, only 16,534 and 420 users, about 50% of users from both interaction graphs, interact with their friends. These results mean that, like other social sharing services, most users usually use the service not for communicating with others, but for making their reading lists and leaving their own impressions. Figure 1 represents the degree distribution of the entire graph and the interaction graph. We omit the extremely large value, which represents the user who is an author and has over 5,000 degrees, in the entire graph of LibraryThing (the graph on left top). In the entire graph, there are extremely popular users who have lots of friends, but most of the users, about 90% have one or zero friends; in the interaction graph, there are also extremely popular users, but the difference from the entire graph is that most users have more than one friend.

The number of users who have one or more friends is 60,317 at LibraryThing and 2,699 at Userstory Book; about 56% and 21% of users make a friend. However, 25,972 users and 1,258 users have only one friend. It implies that the number of users who actively make social relationships is very low, and most users are indifferent to social behavior.

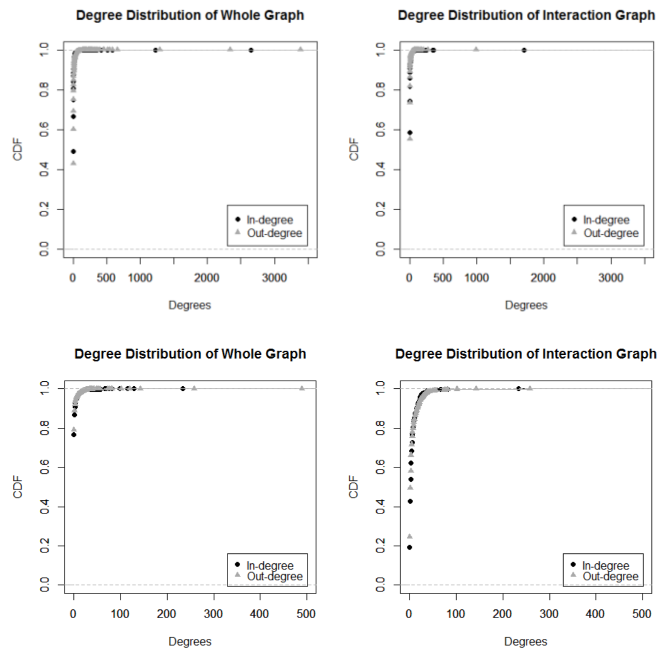


Figure 1. The degree distribution of the entire graph and the interaction graph for LibraryThing (top) and Userstory Book (bottom)

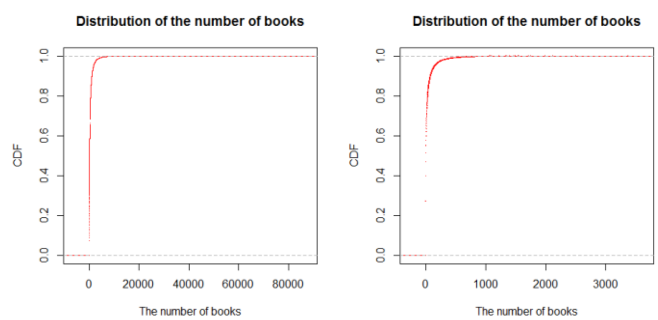


Figure 2. A cumulative distribution of books by how many a user has, for LibraryThing (left) and Userstory Book (right)

Figure 2 shows the distribution of the number of books each user has. There are also extremely active users who have many books in their list, but, in the case of Userstory book, most of the books' status is set to "plans to read" and genre distribution is not skewed. It means the extremely active users are better than the rest in expressing their interest in all genres. According to our estimation, on average, each user has 5.31 friends and 388.25 books at LibraryThing and 1.05 friends and 31.65 books at Userstory Book, including books planned to be read. We also examine the relationships between the number of books and user popularity; the result is in Figure 3. Figure 3 shows that the most active users who have many books are not more popular than general users. The Pearson's

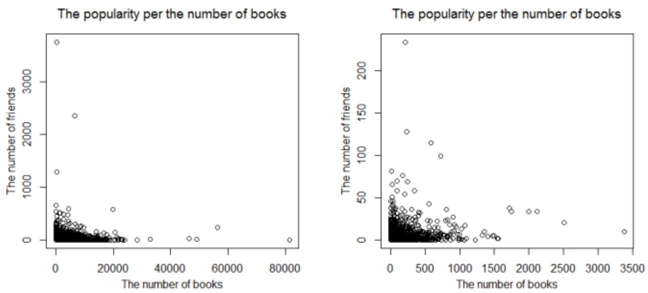


Figure 3. The relation between a user's popularity and the number of books he or she has, for LibraryThing (left) and Userstory Book (right)

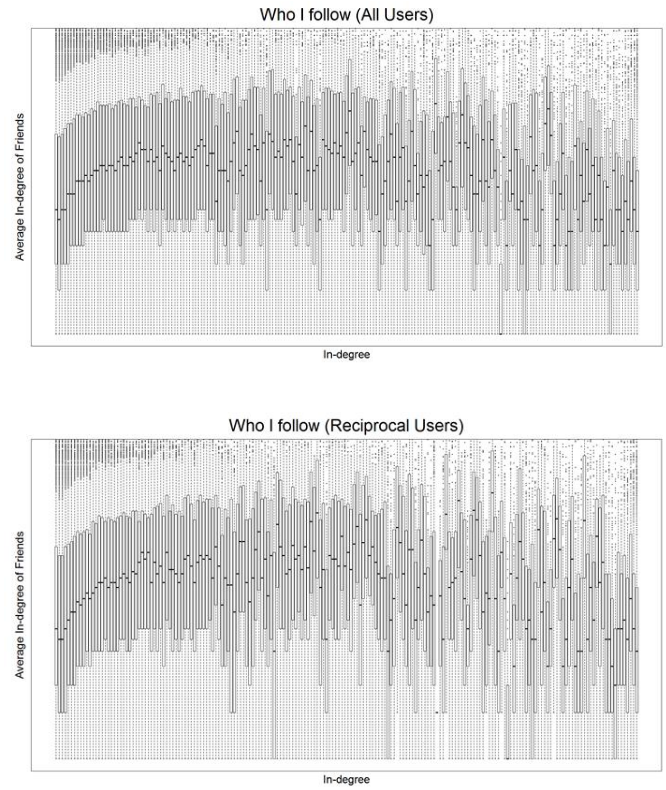
correlation coefficient is 0.154 and 0.276 for LibraryThing and Userstory Book, respectively. Active behaviors like reading many books and expressing one's interest does not affect making relationships and vice versa. This indicates that users present their impressions not for giving information to others, but for their own needs and self-satisfaction. Also, it shows that a social cataloging service is different from a social networking service in that active users have lots of friends.

4 Characteristics of Long Content Service

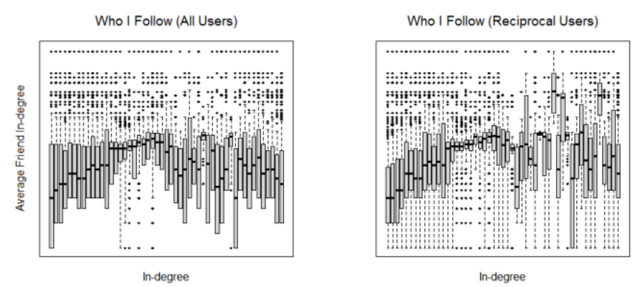
Wattenhofer *et al*. [1] find the differences between a YouTube network and a traditional online social network in terms of assortativity, reciprocity, and homophily. We analyze these features using our dataset and compare them to the result of [1] in order to figure out whether distinctions exist depending on content consumption time.

4.1 Assortativity

Assortativity is the tendency of nodes to connect with other nodes with similar degrees of a certain unit. We examine assortative links based on user popularity. The results are shown in figure 4. The x-axis represents the in-degree of the users, and the y-axis represents the average in-degree of their friends. Figure 4(a) is the result of the LibraryThing dataset and figure 4(b) is the result of the Userstory Book dataset. The plot shows that users from social relationships with others who have a certain amount of in-degrees regardless of the number of in-degrees of themselves, in both datasets. This tendency has nothing to do with the type of the links. The assortativity measurement of the subscription network in YouTube [1] shows that most users in a subscription graph subscribe to a publisher whose popularity is above a certain threshold, and significant differences depend on the type of links; reciprocal users are more assortative than the entire userbase.



(a)



(b)

Figure 4. Similarities between users and their friends' popularities using (a) the entire users and the reciprocal users in the LibraryThing dataset and (b) the entire users of the and the reciprocal users in the Userstory Book dataset. There is no difference in assoravity between the entire users and the reciprocal users in both datasets.

4.2 Reciprocity

In our dataset, most links are reciprocal; 95.93% of links in the LibraryThing dataset and 73.32% of the Userstory Book dataset are bidirectional. That is, the number of the users who join in interacting with each other is less, but the level of reciprocity is very high. This implies that the relationships are usually superficial and tend to be passive engagements [14]. Most of the users keep up with friends by reading news, reviews, ratings, and such without actual communications. They do not try to develop deeper connections.

Our results are quite large compared to the reciprocities of other directed social networks such as YouTube at 25.42% [1], Filckr at 68% [15], and the sampling data of CiteULike which have only 93 reciprocal links out of 11,295 unilateral links [9]. The details of the results are in Table 2.

Table 2. The proportion of reciprocity

	Library Thing	Userstory Book
The number of links	60,317	2,699
Reciprocal links	57,864	1,979
	95.93%	73.32%

4.3 Homophily

We use genres of the books a user has to measure homophily among users. For this analysis, we only use the Userstory Book dataset because LibraryThing does not provide any information on genres. The books in a user's list cover a wide range of genres. Because there is more than one genre, we assume that the most read genre in the user list represents the user's primary interest. Prior to the analysis, we classified 29 different genres based on how online bookstores generally categorize their genres, because the genre classification in Userstory Book is too specific.

We measure the level of homophily in two graphs; the entire graph and the interaction graph. Regardless of the type of graph, about 50% of users are interested in the same genre as their friends: 49.24% from the entire graph and 52.25% from the interaction graph. This indicates that the social cataloging services are more homophilous than the YouTube network, which is on average 26.58% and 27.46% respectively [1]. We expect to get better results if the books are classified into fewer genres.

5 Influential Features in Long Content Services

Although the social content sharing services that provide longer content have low communication ratios, a user's content selection is not independent from that of others; users cannot be completely isolated from social influence because of the various social features provided by the services. We measure features that affect users' item selection.

We begin by examining the relational features such as friend and interaction networks. We assume that a user's friend influences the user the most because of the high reciprocity ratio. However, table 3 shows that 48,674 users and 1,759 users of both services share the same books with their friends, and about 20% of the books in a user's reading list show up in the user's friends' reading lists. In the case of interactional relationships, the percentage of user-friend reading list overlap is less than the results found in friend relationships. This is because half of the users who leave messages on others' pages have no relationships with the recipient, and the messages have little to do with books in particularly. The diversity of the books also affects this results.

Table 3. The result of social influence from friends: each column denotes (a) how many books a user has in common with his or her friends, (b) how many friends have one or more common books with each user, (c) the percentage of overlap between a user's book list and that of his or her friends, and (d) the number of users who have the same books with the list of their friends.

		(a)	(b)	(c)	(d)
Library Thing	AVG	127.36	4.94	24.33%	48,674
	SD	377.82	17.91	22.70%	
Userstory Book	AVG	17.31	4.44	22.80%	1,759
	SD	49.05	9.55	22.33%	

Table 4. Social influence of users who leave messages in others' pages: columns are the same as in table 3.

		(a)	(b)	(c)	(d)
Library Thing	AVG	108.07	2.96	16.60%	24,506
	SD	321.55	7.64	18.46%	
Userstory Book	AVG	10.05	1.86	13.83%	453
	SD	20.75	2.46	16.74%	

Next, we measure how similar a user is to his or her friends in terms of the genre composition of their reading lists. For this analysis, we only use the Userstory Book dataset and the 29 genres we classified in the homophily experiment because of the absence of genre information in LibraryThing. We consider that each user has an n-dimensional vector $\vec{v}(u_i)$,

$$\vec{v}(u_i) = (g_1, g_2, \cdots, g_n) \quad (1)$$

where n is the number of genres and g_k is the number of books belonging to the k^{th} genre the user u_i has. We set $n = 29$. For more meaningful results, we only focus on users who have ten or more books. The genre similarity between users are measured using cosine similarity [16],

$$GenreSim(\vec{u},\vec{v}) = Cosine(\vec{u},\vec{v}) = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} \quad (2)$$

Figure 5 shows the genre similarity. We observe that the genre similarity among users with social relationships is higher than that of all possible pairs of users. Among social relationships, the similarity of those with reciprocal relationships is a little higher. That is, users may consider common interests when choosing friends, and these relationships in turn can influence their item selection. As in the homophily experiment, we expect to get better results if we deal with fewer genres.

Shared recent activities of others' are helpful for users in choosing new books. Books can be registered in the reading list either before or after the reading. In the latter case, there is a time difference between beginning the book and registering the books after having completely read it. We assume that the average period of reading a book is within three days[1], and examine whether there are other users who have already registered the book before within the three days of another user's registration of the same book. If such cases happen quite often, we can consider that a user may be influenced by others' recent activities; we guess that there is a high possibility the user looks at the other's recently registered book in the news feed. Figure 6 shows results from the entire Userstory Book dataset. We also only use the Userstory Book dataset because they provide all users' activities. We find that about 80% of the books are not overlapped, a user's registration of a certain book does not happen during the three days after that same book is registered by any other user. This means people may see a book in other users' recent activities, but they do not select that book as their next item. We conclude that recent activity is not an influential feature in passive engagement.

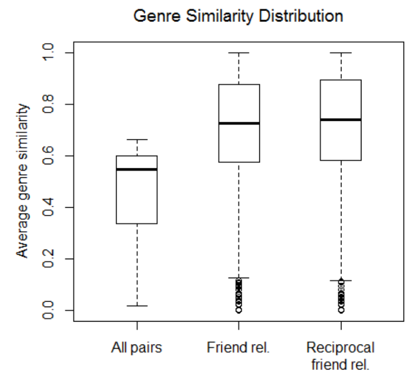


Figure 5. The distribution of genre similarity

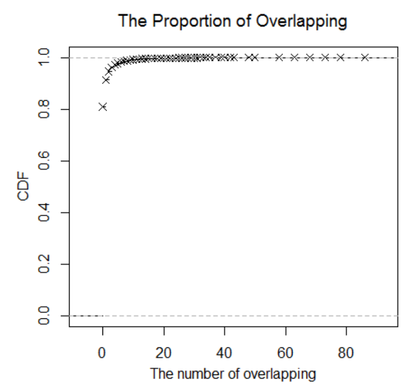


Figure 6. The proportion of the book overlapping between a user's reading list and others' recent activities

6 Conclusion

In this paper, we classified social content sharing services based on the consumption time of the target content. To figure out the characteristics of networks that have long content, we used two social cataloging services, LibraryThing and Userstory Book. The dataset of Userstory Book is quite smaller than the dataset of LibraryThing, but it also meant we were able to work on a comprehensive dataset of a social cataloging service. This is good for figuring out the general characteristics of long content services. We analyzed and compared the datasets to the YouTube network [1], a representative short content service, in terms of assortativity, reciprocity, and homophily. According to results, our datasets have low assortativity, but the level of reciprocity and homophily is higher than those of the YouTube network. We also study the social features in both datasets that affect users' item selection. From our results, we find that interest similarity is an important factor in determining users' item selection. We found that the genre of books also plays an important role in forming user relationships and selecting new items. For future work, we hope to focus deeper into social features of networks with long content such as the effect of communities and more detailed statistical information.

7 Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIP) (No. 20110030812), by ETRI R&D Program ("Development of Big Data Platform for Dual Mode Batch-Query Analytics, 14ZS1400") funded by the Government of Korea, and by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIP) (No. 20110017480)

[1] We searched polls about the average time it takes to read a book. In these polls, more than half of people answered that they usually take 1 to 3 days to read a book. The sources are as follows: http://www.goodreads.com/poll/show/45995-how-long -does-it-take-you-to-finish-an-average-size-book-approx and http://dearauthor.com/features/poll-misc/poll-how-long-does-it-take-to-read-a-book/#ViewPollResults

8 References

[1] M. Wattenhofer, R. Wattenhofer, and Z. Zhu, "The YouTube Social Network," in *Proceedings of International AAAI Conference on Weblogs and Social Media*, 2012, pp. 354–361.

[2] M. Papagelis and V. Murdock, "Individual Behavior and Social Influence in Online Social Systems," *Proc. 22nd ACM Conf. Hypertext hypermedia - HT '11*, pp. 241–250, 2011.

[3] L. Rashotte, "Social Influence," *blackwell Encycl. Soc. Psychol.*, vol. 26, no. 3, pp. 562–563, 2007.

[4] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Bus. Horiz.*, vol. 53, no. 1, pp. 59–68, Jan. 2010.

[5] N. Cook and D. Tapscott, *Enterprise 2.0*. 2008, p. 180.

[6] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," *Proc. 7th ACM SIGCOMM Conf. Internet Meas. - IMC '07*, p. 29, 2007.

[7] L.-M. Aiello, A. Barrat, C. Cattuto, G. Ruffo, and R. Schifanella, "Link Creation and Profile Alignment in the aNobii Social Network," *Soc. Comput. (SocialCom), 2010 IEEE Second Int. Conf.*, 2010.

[8] M. Tang, Y. Ke, and Y. Sie, "The Estimation of aNobii Users ' Reading Diversity Using Book Co-ownership Data : A Social Analytical Approach," in *The 5th International Conference on Social Informatics (SocInfo)*, 2013, pp. 274–283.

[9] D. H. Lee and P. Brusilovsky, "Social Networks and Interest Similarity: The Case of CiteULike," in *Proceedings of the 21st ACM conference on Hypertext and hypermedia HT 10*, 2010, pp. 151–155.

[10] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri, "Feedback effects between similarity and social influence in online communities," *Proceeding 14th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. KDD 08*, vol. 363, no. 3, p. 160, 2008.

[11] C. A. Yeung, "Capturing Implicit User Influence in Online Social Sharing Categories and Subject Descriptors," *Proc. 21st ACM Conf. Hypertext hypermedia HT 10*, pp. 245–254, 2010.

[12] J. Tang, J. Sun, C. Wang, and Z. Yang, "Social influence analysis in large-scale networks," *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '09*, p. 807, 2009.

[13] Y. Caiv and Y. Chen, "Mining influential bloggers: From general to domain specific," in *KnowledgeBased and Intelligent Information and Engineering Systems*, 2009, vol. 5712 LNAI, pp. 447–454.

[14] C. Marlow, L. Byron, T. Lento, and I. Rosenn, "Maintained Relationships on Facebook," 2009. [Online]. Available: http://overstated.net/2009/03/09/maintained-relationships-on-facebook.

[15] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," *Proc. 18th Int. Conf. World wide web - WWW '09*, p. 721, 2009.

[16] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Addison-Wesley, 2005.

Correlation and Regression Models to Assess the Usability of the Web Pages

Saad Subair[1] and Hussah AlEisa[2]
[1]Princess Nourah bint Abdulrahman University, Riyadh, KSA
[2]King Abdulaziz City for Science and Technology, Riyadh, KSA

Abstract *In this research work, the usability of the web pages is assessed using correlation and regression models. Researchers argue that many usability results and recommendations lack empirical and experimental data. Universities web pages are chosen as subjects for this work. Series of experiments have been conducted to investigate into the usability and design of the universities web pages. Prototype web pages have been developed according to the structured methodologies of web pages design and development. Four universities web pages were evaluated together with the prototype web pages using a questionnaire which was designed according to the Human Computer Interactions (HCI) concepts. The data collected is analysed using correlations and regression models. The correlation analysis showed that there are significant positive and negative correlations between many items. The regression analysis revealed that the most significant factors (items) that contributed to the best model of the universities web pages design and usability were: multimedia in the web pages, the web pages icons (alone) organisation and design, and graphics attractiveness. The results showed some of the limitations of some heuristics used in conventional interface systems design and proposed some additional heuristics in web pages design and usability.*

Keywords: Web Pages Usability, Human Computer Interaction (HCI), Correlation Model, Regression Model.

1. Introduction

User interface is a system layer through which the computer and users communicate [1, 2]. It is an essential part of Human Computer Interaction (HCI). Dealing with or achieving goals using the interface is known as the usability of the interface. Usability of a system is how easy the system to use and how easy and efficient to perform tasks [3, 4]. In a survey conducted in the last few years, software development devoted almost half (48 %) of the code to the user interface [5, 6]. It is important to devote considerable time, effort, and cost for user interface design and its usability. Usability of the interface is an important issue. Nielsen [4, 7] conducted a usability test for software used by a company. The result concluded that there was a list of 130 usability problem. He concluded that the software design was sound and most of the usability problems were simple enough to fix. As far as usability benefits are concerned, cost saving is considered one of the most important benefits of usability. Many researchers [8, 9, 10, 11, 12, 13] presented several examples of significant cost saving due to usability studies. Usable system lead to lower training costs, decreased requirement for field support, and increased user satisfaction [14, 15]. The interfaces used in the web pages are graphical user interfaces that utilize graphics, colours, and icons. Researches showed that there is still a big gap between the researches of the HCI and hypertext systems, essentially the web [16]. Shneiderman [17] argued that many researchers' experience lack of empirical data to validate or solidify their conclusions. Shneiderman [17] reported that web sites can be well categorized by the originator's identity such as individual group, university, corporation; non-profit organization. He concluded that information about users can guide web designers to a better design. However, the problems that motivated this research are:

I. There were no empirical studies that categorized the users of the universities web pages; study their preference, and the problems they are facing when using these web pages.
II. To specify the relative importance of some items or parameters of interface design and usability.

The main objective of this study is to investigate into the relations between the items of web pages design and their usability and to point out which items or variables contribute significantly to web pages design of the universities. Questionnaires were used to evaluate the usability of interfaces. However, researchers found problems and weaknesses in the questionnaires developed in the past [18, 19]. The problems varied from non-represented population sample, to lack of validation and low reliabilities of the questionnaires. User acceptance or subjective satisfaction of a system is a critical measure of the system success. Chin et al [20] developed a measurement instrument which measures the user's subjective rating of the human computer interface called the Questionnaire for User Interface Satisfaction (QUIS). Improvements to the questionnaires for evaluating computer systems and interfaces were suggested and introduced by several researchers [21, 22].

2. Methodology

In this study, the problems and the objectives of this work which are stated earlier were attempted to be solved through:

1- Designing and developing a prototype user interface for university web pages using recent web pages development methodologies.
2- Evaluating the prototype and other universities web pages usability in a comparative approach through a questionnaire.
3- Using correlation analysis to investigate the relationship and association between items or variables of universities web pages design.
4- Using regression analysis to investigate the causality relationships to find out the most significant independent variables (items) that form the best model or models for the universities web pages design.

The experiment conducted was a comparative evaluation of the usability variables of these web pages. A sample of 60 students was given a questionnaire to respond to. Data of the questionnaire were collected, organized, and saved in a flat file format to be ready to be transferred and manipulated by other packages and programs. SAS and SPSS statistical packages were used to statistically analyze the data. The questionnaire was designed using a rating scale ascending from 1 to 5 which was designed to be administered at controlled experimental conditions. The questions were designed such that each question represents an item of design or heuristic of the interface for the universities web pages. Each item was considered as a variable contributing to the web pages design. In the big experiment of this research, originally several items or variables of web pages design and usability were designed as questions in the questionnaire. In this study only 14 items were considered for the correlation and regression models. They are:

1- Color amount
2- Contrast against the background
3- Graphics information provision
4- Color and Graphics feel of Consistency
5- Navigation
6- Location tracking inside the web pages
7- Hyperlinks (hotspot) semantics
8- Forms organization and helpfulness
9- Animated pictures effect on web pages look
10- Text (alone) organization and design
11- Icons (alone) organization and design
12- Icons and text organization and design
13- Multimedia preference
14- Universities web pages ranking

2.1 Correlation Analysis

Simple correlation between the 14 items of universities web pages design were calculated from the original data. The correlation coefficient (r) between two characters (X and Y) is usually expressed as follows:

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Where

x = x - X and

y = y - Y and

x = Independent variable (item)
X is the mean of character x.
y = Independent variable (item)
Y is the mean of character y.

2.2 Regression Analysis

Simple and multiple linear regressions of some items (independent variables) of the ranking of the university web pages item (dependent variable) were estimated. The formula for multiple linear regression is given as follows:

$$Y = a + b_1 X_1 + b_2 X_2 + \cdots b_n X_n$$

Where:

Y = dependent variable

X_{1-n} = independent variables

a = intercept of the regression line on the Y axis

b_{1-n} = linear regression coefficients

n = Number of independent variables

Backward, forward, and stepwise selection regression analysis had also been attempted. In stepwise selection variables are added (or removed) one by one into (from) the model. The F-statistic is calculated after a variable is added into (or removed from) the model. A variable that did not produce significant F-value at a predefined probability level is then deleted from (or added to) the model. In forward selection, variables already in the model do not necessary stay there and the process ends when none of the variables outside the model is significantly contributing to the model at the specified probability level [23].

3. Results and Discussion

The descriptive statistics of the items (variables) designed to investigate the user interface design and usability in this study is calculated. Low variances in the results show continuous data and hence suggest normal distribution of the data. In this paper we will discuss the correlation and regression analysis only.

3.1 Correlation Analysis

The correlation analysis study is used to explain the relationship between the variables of the web pages design and usability. The objective of this model is investigate whatever two variables (items) are related positively or negatively to each other. Table 1 shows the correlation coefficients between the 14 variables studied. The web pages color amount has no significant correlation with any of the variables of the web pages studied. The web pages contrast against the background showed significant positive correlation with the web pages icons and text organization, color and graphics consistency, navigation, location tracking, and forms organization and helpfulness. However, the contrast against the background showed no significant correlation with the web pages animated pictures effect and icons (alone) organization. The web pages graphics information provision has significant positive correlation with all the other variables studied. Specifically, it showed high positive correlation with the web pages hyperlinks semantics (Table 1). Table 1 shows that web pages color and graphics feel of consistency has positive significant correlation with most of the variables studied. However, it has no significant correlation with the text (alone) and icons (alone) organization but has high positive significant correlation with the icons and text together organization and helpfulness. Navigation of the web pages has significant positive correlation with all the variables studied except the web pages icons (alone) organization and the multimedia preferences in the universities web pages. Location tracking and hyperlinks semantics showed significant positive correlation (separately) with almost all the other variables studied except the ranking of the universities web pages. This result indicated that the ranking of the web pages of the universities did not rely much on these two variables. The web pages forms organization has significant positive correlation with the rest of the variables studied. Forms organization associated significantly with the ranking of the universities web pages. The animated pictures effect on the web pages showed significant positive correlation with the web pages text (alone), icons (alone), and icons and text (together) organization and helpfulness.

However, animated pictures had no significant correlation with the ranking of the web pages. Text (alone) organization has significant positive correlation with the icons alone organization and multimedia preferences in the web pages of the universities. Text (alone) has no significant correlation with the ranking of the universities web pages. Icons (alone) organization has significant positive correlation with the icons and text (together) organization, multimedia preference, and the web pages ranking. Icons and text (together) organization and multimedia preference had significant positive correlation (separately) with the ranking of the universities pages.

3.2 Regression Analysis

As mentioned earlier, regression analysis is used to understand how the value of the dependent variable changes when any one of the independent variables is varied. That means regression analysis estimates the conditional anticipation of the dependent variable given the independent variables. Regression analysis is also used to understand which among the independent variables are related to the dependent variable; and to explore the types of these relationships. In some cases, regression analysis can be used to deduce causal relationships between the independent and dependent variables. Here we use the regression model to deduce which independent variable or variables contribute significantly to the dependent variable web pages ranking. The multiple linear regression results are shown in Table 2. The ANOVA (analysis of variance) table for regression shows that the model was significant at the 0.01 probability level, indicating that the contribution of independent variables coefficients were not equal to zero (alternate hypothesis). Simply, it says that there is a contribution from some variables to the model at a certain level of probability (Table 2). The table clearly shows that multimedia preference in the universities web pages contributed highly and significantly to the model. However, some regression analysis procedures were used, including, forward selection, backward elimination, and stepwise selection. Maximum coefficient of determination (R^2) was achieved by including specific number of independent variables. The forward selection and backward elimination adds and eliminates (respectively) independent variables to (from) the model regardless of what variables are already outside or existing in the model (tables not shown). Stepwise selection re-examines at every stage the regression of the variables incorporated into the model in previous stages. A variable which might have been the best single variable to enter the model at an early stage, might at later stage, be superfluous because of the relationships between it and other variables now in the regression model [23].

Table 1: Correlation coefficients of the 14 items (variables) of the universities web pages design and usability

	Colour amount	Contrast against the back-ground	Graphics information provision	Colour and graphics feel of consistency	Navigation	Location Tracking	Hyperlink Semantics	Forms organization	Animated pictures effect	Text (alone) organization	Icons (alone) organization	Icons and text organization	Multimedia preference
Contrast against the background	0.05												
Graphics information provision	0.11	0.28**											
Colour and graphics feel of consistency	0.01	0.33**	0.27**										
Navigation	0.09	0.28**	-	0.14*									
Location Tracking	0.05	0.28**	0.25**	0.26**	0.41**								
Hypelinks Semantics	0.03	0.21**	0.34**	0.21**	0.33**	0.40**							
Forms organization	-0.04	0.27**	0.23**	0.30**	0.33**	0.41**	0.37**						
Animated pictures effect	0.00	0.13	-	0.20**	0.10	0.31**	0.22**	0.31**					
Text (alone) organization	0.01	0.20**	0.16**	0.11	0.22**	0.29**	0.26**	0.41**	0.18**				
Icons (alone) organization	0.07	0.09	0.21**	0.11	0.11	-	0.27**	0.21**	0.24**	0.17*			
Icons and text organization	0.06	0.35**	0.28**	0.33**	0.32**	0.44**	0.26**	0.26**	0.35**	-	0.27**		
Multimedia preference	0.13	0.21**	0.20**	0.20**	0.11	0.18**	0.20**	0.29**	-	0.35**	0.24**	-	
Universities web pages ranking	0.11	0.15*	0.15*	0.19**	-	0.03	0.10	0.17*	0.13	0.09	0.18*	0.18**	0.27**

** : Significant at 0.05 level of probability.
* : Significant at 0.10 level of probability.

Table 2: Linear regression analysis of items (variables) of web pages design and usability (web pages ranking is the dependent variable)

Source	Degrees of Freedom	Sum of Square	Mean Square	F	p > F
Regression	13	70.00086	5.384682	7.26	0.0001
Error	243	180.1159	0.741218		
Total	256	250.1167			

Variable	Parameter Estimate	Standard Error	Sum of Squares	F	p >F
Intercept	1.061532	0.440591	4.302695	5.80	0.0167
Color amount	0.027230	0.075147	0.097328	0.13	0.7174
Contrast against the background	-0.030090	0.059111	0.192028	0.26	0.6112
Graphics information provision	0.068228	0.074656	0.619078	0.84	0.3617
Color and graphics feel of consistency	-0.030560	0.071732	0.134551	0.18	0.6704
General navigation	0.069340	0.068237	0.765360	1.03	0.3106
Location Tracking	-0.064310	0.069323	0.637798	0.86	0.3545
Hyperlinks (hotspots) semantics	0.121541	0.073557	2.023700	2.73	0.0998
Forms organization and helpfulness	0.026652	0.070631	0.105537	0.14	0.7063
Animated pictures	-0.055670	0.089501	0.286755	0.39	0.5345
Text (alone) organization and design	-0.109380	0.064121	2.157001	2.91	0.0893
Icons (alone) organization and design	0.200612	0.057933	8.888093	11.99	0.0006
Icons and text organization and design	-0.018650	0.060757	0.069868	0.09	0.7591
Multimedia preference	0.416697	0.073722	23.68064	31.95	0.0001

Table 3 further shows a summary of stepwise selection regression procedure model which included only four variables that were significantly contributed to the model at the 0.15 probability level. That is to say, universities web pages multimedia preference, icons (alone) organization, text (alone) organization, and hyperlinks (hotspots) semantics were the four most independent variables contributing to the model significantly.

Coefficient of determination (R^2) measures the proportion of total sum of squares of the variables that is explained by the regression line (table not shown). It is a measure of how closely the points (observation) fit the least square line. Consequently, the line that has maximum R^2 represents the best fitting line [24]. This might indicate the ranking importance of an independent variable as a single or in combination with other independent variables to universities web pages ranking.

Table 3: Summary of stepwise procedure for dependent variable universities web pages ranking [a]

Step	Variable Entered / Removed	Number in	Partial R^2	Model R^2	P > F
1	Multimedia preference	1	0.2132	0.2132	0.0001
2	Icons (alone) organization and design	2	0.0401	0.2534	0.0003
3	Text (alone) organization and design	3	0.0073	0.2607	0.1151
4	Hyperlinks semantic	4	0.0094	0.2700	0.0736

a : All variables in the model are significant at 0.15 probability level.

However, the results showed that the universities web pages location tracking, hyperlinks semantics, text (alone) organization, icons (alone) organization, and design, and multimedia preferences were among the first five variables (items) that contributed significantly to the universities web pages ranking model. The limitation to R^2 is that the addition of an irrelevant independent variable to the regression model will increase R^2 even though the irrelevant variable is not related to the other variables in the model [24]. The assumption of this model is that all the variables (items) have been estimated precisely. This might not always be true. Overall, the results suggest that, some independent variables significantly contributed to the regression model and others did not. For a single independent variable, multimedia preference of the universities web pages gave the most significant contribution to the dependent variable universities web pages ranking. For two independent variables, icons (alone) organization and design and multimedia preference of the universities web pages gave the most significant contribution to the dependent variable universities web pages ranking. For three independent variables, the universities web pages text (alone) organization and design, icons (alone) organization and design, and multimedia preference gave the most significant contribution to the dependent variable universities web pages ranking.

4. Conclusions

Unveiled were important conclusions and findings concerning some items (variables) of the web pages design and usability. These conclusions and findings showed the power of the correlation and regression models to figure out the importance and ranking of the variables that contribute significantly to the usability of the web pages. These models can be applied and extended to any other variables or (items) of web pages design to assess their usability importance.

5. Acknowledgments

The authors would like to thank Princess Nourah bint Abdulrahman University, Riyadh, KSA and King Abdulaziz City for Science and Technology, Riyadh, KSA for supporting this research, also they would like to thank all the students responded to the questionnaire.

6. References

[1] Jacko, Julie A . (2012)The human-computer interaction handbook: fundamentals, evolving technologies, and emerging applications p.1518. CRC Press.

[2] Nielsen, J. (1993) Usability Engineering. Academic Press, Inc. Harcourt Brace & Company, Publishers. San Diego USA.

[3] Shneiderman B, Kent L. Norman, Catherine P, Benjamin B. Bederson, Druin A, Golbeck J (2013): “30 years at the University of Maryland's human-computer interaction lab (HCIL)”. Interactions 20(5): 50-57.

[4] Nielsen, J. (2013). Usability 101: Introduction to usability. Jakob Nielsen’s Alertbox, 25.

[5] Myers, B. A., and Rosson, M. B. (1992). "Survey on user interface programming". Proc. ACM CHI'92 Conf. (Monterey, CA, 3-7 May), 195-202.

[6] Norman, D. A. (2010). Natural User Interfaces Are Not Natural. Interactions, 17, No 3

[7] Nielsen, J. (l990b). "Big paybacks from 'discount' usability engineering". IEEE Software 7,3 (May), 107-108.

[8] Davis, P., & Shipman, F. (2011). Learning usability assessment models for web sites. Proceedings of the 16th international conference on intelligent user interfaces (pp. 195–204).

[9] Albert, W., Gribbons, W., & Almadas, J. (2009). PreConscious Assessment of Trust: A Case Study of Financial and Health Care Web Sites. Human Factors and Ergonomics Society Annual Meeting Proceedings, 53, 449453.

[10] Karlin, J. E., and Klemmer, E. T. (1989). An interview. In Klemmer, E. T. (Ed.), Ergonomics: Harness the Power of Human Factors in Your Business. Ablex, Norwood, NJ. 197-201.

[11] Fisher, R, and SIess, D. (1990). "Information design methods and productivity in the insurance industry". Information Design Journal 6, 2, 103-129.

[12] Bailey, W. A., Knox, S. T., and Lynch, E. R (1988). "Effects of interface design upon user productivity". Proc. ACM CHI'88 Conf. (Washington, DC, 15-19 May), 207-212.

[13] Karat, C. (1990). "Cost-benefit analysis of iterative usability testing". Proc. IFIP INTEPACT'90 Third IntI. Conf. Human-Computer Interaction (Cambridge, U.K., 27-31 August), 351-356.

[14] Landseadel, P. (1994). "Methodologies in User Interface Design". IEEE Electronic Library. (No. 0-7803-1910-9/94).

[15] Norman, D. and Nielsen, J (2013). 10 Heuristics for User Interface Design. Retrieved from http://www.nngroup.com/articles/ten-usability-heuristics/

[16] Buckingham, S. and McKnight, C. (1997) “Word wide web usability: introduction to this special issue”. Int. J. Human-Computer Studies, 47, (1), 1-4.

[17] Shneiderman, B. (1997) "Designing information-abundant web sites: issues and recommendations". Int. J. Human-Computer Studies, 47, (1), 5-30.

[18] Bailey, J. E. and Person, S. W. (1983) “ Development of a tool for measuring and analysing computer user satisfaction”. Management Science, 29, 5, 530-545.

[19] Larcker, D. F. and Lessig, V. P. (1980) "Perceived usefulness of information: A psychometric examination". Decision Science, 11, 1, 121-134.

[20] Chin, J. P., Virginia, A. D., and Norman, K. L. (1988) “Development of an instrument measuring user satisfaction of the Human-Computer interface” Proceedings of the CHI’88, ACM. SIGCHI Bulletin. Special issue ACM order No. ACM-0-89791-265-9/88/0004/0213.

[21] Rushinek, A. and Rushinek, S. F. (1986) "What makes user happy?". Communication of the ACM, 29, 7, 594-598.

[22] Subair, S (2014) "The Usability Engineering of the Universities Web Pages User Interfaces: A Pragmatic Approach". Princess Nourah bint Abdulrahamn University Printing Press. Riyadh, KSA.(Under Press)

[23] Drapper, N. R. and H. Smith. (1981) Applied Regression Analysis. John Wiley & Sons Inc. N.Y. USA.

[24] Retherfor, R. D. and M. K. Choe (1993) Statistical Models For Causal Analysis. John Wiley & Sons, Inc.

DESIGN AND DEVELOPMENT OF A SYSTEM FOR DATA CAPTURE IN A MONITORING AND EVALUATION ENVIRONMENT

Patrick Otoo Bobbie*
Faculty of Informatics
Ghana Technology University College
Tesano – Accra, Ghana
pbobbie@gtuc.edu.gh

Joshua Ronald Opoku Nsiah
Faculty of Informatics
Ghana Technology University College
Tesano – Accra, Ghana
nsiahopoku@yahoo.com

**Also with Southern Polytechnic State Univ, Marietta, GA, USA*

Abstract

Monitoring and evaluation (M&E) is an extremely complex, multidisciplinary and skill intensive endeavor. Government-wide M&E is even more so because it requires detailed knowledge across and within sectors. New technologies are changing the nature of monitoring and evaluation in the world today. Information and communication technology (ICT) tools such as computers, mobile phones and tablets, together with applications/software systems which allow users to upload data to storage facilities in real-time are being utilized. The use of mobile computing tools and the Internet in this research provides an efficient and effective way of performing Government-wide M&E functions. The system provides functionality for data capture and analyses in an M&E environment. It utilizes the internet to reach out to all stakeholders in the M&E processes. Furthermore it uses the Short Message Service (SMS) for its real-time impact assessment functions. The system was developed using open source development tools and languages. A working prototype version of the system, has been deployed successfully on an experimental basis.

Keywords: *SMS/ICT, Monitoring, Evaluation*

1 Introduction and Background

One major challenge faced by African governments is to be more effective. Monitoring and Evaluation (M&E) processes or procedures can assist the public sector in evaluating its performance and identifying the factors which contribute to its service delivery outcomes.

M&E is uniquely oriented towards providing its users with the ability to draw causal connections between the choice of policy priorities, the resourcing of those policy objectives, the programs designed to implement them, the projects or services actually delivered and their ultimate impact on communities. In this context, a government-wide M&E policy helps to provide an evidential basis for public resource allocation decisions and helps identify how challenges should be addressed and successes replicated.

In most cases, African governments are at varying points of establishing government-wide M&E systems. Countries such as Ghana, South Africa, Uganda and Kenya are using M&E to assess its progress against national development plans.

1.1 The Problem

In Ghana, the Policy Unit at the Presidency has the mandate to effectively manage government policies, programs and projects towards the achievement of the government's vision. One way of carrying out this mandate is through the monitoring and evaluation of all government policies, programs and projects. It has the mandate to monitor and evaluate the performance of 23 ministries, 10 regional coordinating councils, 275 metropolitan, municipal and district assemblies and over 290 departments and agencies. This means that on a daily basis, the policy unit at the presidency needs to monitor over 598 government agencies which are dispersed throughout the country.

However currently, feedback from the various government agencies are sent to the policy unit on a monthly basis, the feedback is paper-based making them prone to error, difficult to conduct on a large scale and high in transaction cost. Furthermore, there is no correspondence with the direct beneficiaries of these policies, programs and projects in order to access the impact of those activities. There are also delays in responses from the various respondents, due to transportation issues. Sometimes posted responses fail to arrive on time or they never even arrive at all. All the issues discussed above represent challenges faced by the policy unit in the execution of their M&E functions.

This research seeks to leverage the use of ICT to solve M&E challenges as outlined above, and as a case study, to deal with issues faced by the Policy Unit at the Presidency in Ghana. An ICT platform has been designed and implemented for M&E data collection and monitoring by the Policy Unit to track the implementation of government policies, programs, and projects by all government agencies. This prototype is web-based, with mobility and database technologies for data collection and storage as well as provisioning of real-time information about the status of policies, programs, and projects to all stakeholders (i.e. the Presidency, beneficiaries, donors, NGO's and partners).

This system also incorporates the use of SMS technology to enhance the project impact assessment functions of the policy unit. It is expected that this system would provide the policy unit with real-time data collection and data analysis functionalities. The M&E system includes various reporting functionalities as well as a simple Executive Information System (EIS) that provides the policy unit with the real-time status of project implementations.

2 Innovation in M&E data collection

2.1 Monitoring and Evaluation

Although the two key functions in the term "monitoring and evaluation" are often convoluted, monitoring and evaluation are, in fact, two distinct sets of organizational activities, related but not identical.

Monitoring and Evaluation (M&E) of development activities provide government officials, development managers, and civil society with better means for learning from past experience, improving service delivery, planning and allocating resources, and demonstrating results as part of accountability to key stakeholders (World Bank, 2004). Monitoring is the systematic collection and analysis of information as a project progresses. It is aimed at improving the efficiency and effectiveness of a project or organization. Evaluation is the comparison of actual project impacts against the agreed strategic plans. It looks at what tasks are to be completed, and how the tasks were accomplished.

2.2 Innovation in Monitoring and Evaluation

According to Raftree (2013b), what ICT is doing for M&E is really broadening it out and allowing more people to participate in it. It also allows project managers to analyze data better and to make decisions more quickly. New ICT tools can make monitoring faster and more accurate. Using mobile phones can also help in reaching out to a wider group, and get feedback from the communities. Some of the new ICT-enabled visualizations tools, including maps, graphs and charts, make it easier to analyze collected data with feedback and sharing of outcomes with the communities.

It is evident that ICT is an important innovative tool to aid in an efficient and effective M&E processes. Its implementation can be seen in diverse sectors to provide solution to budding problems. According to Zanamwe and Okunoye (2013), Information and Communication Technologies play an important role in significantly mitigating climate change and these technologies can be used in both developed and developing countries even though developing nations lack the much needed information and communication technology infrastructure when it comes to mitigating climate change.

The UNDP (2013) report on innovations in monitoring and evaluating results provide many good examples of ICT technologies that are providing innovative ways of performing monitoring and evaluation in various fields of endeavors. Despite all this, it is important to note that ICT for M&E is still emerging and there are several challenges in the out-years.

2.3 ICT as a tool for data collection

Data collection has always been an integral part of monitoring and evaluation (M&E). Collecting data on policies, programmes and projects by government, helps the government to efficiently monitor and evaluate the impact of these activities on its populace. One way to facilitate as well as reduce the cost of data collection, is the use of ICT tools.

Examples of ICT-based technologies which aid in data collection include: Mobile phones (Short Messaging Service), Personal Digital Assistants (PDA's), Web Technologies, Audio Computer-assisted Self-Interviewing, Mapping & Geographic Information Systems, Photo/Video Monitoring, and Social Media Channels. For the purpose of this research, two of these ICT tools (i.e. Web and SMS) are discussed in the following.

2.3.1 Web as a tool for data collection

Although the World-Wide Web was initially conceived as a vehicle for delivering and viewing documents, its focus has gradually shifted from documents to applications. Facilities such as Javascript, the Document Object Model (DOM), and Ajax have made it possible to offer sophisticated interactive applications over the Web.

Web-based applications run on a web application server and access data on an enterprise information system, such as a MYSQL database server. The components of web-based applications are spread across multiple tiers, or layers. In general, the user interface is on the first or top tier, the application programs are on the middle tier, and the data sources that are available to the application programs are on the enterprise information system tier. Developing web-based applications across a multi-tiered architecture is referred to as server-side programming. Writing server-side programs is complicated and requires a detailed understanding of web server interfaces. Fortunately, application servers, such as Apache Geronimo Server, are available to simplify this task. Each of these application servers defines a development environment for web applications and provides a runtime environment in which the web applications can execute. The application server code, which provides the runtime environment, supports the appropriate interface for interacting with the web server.

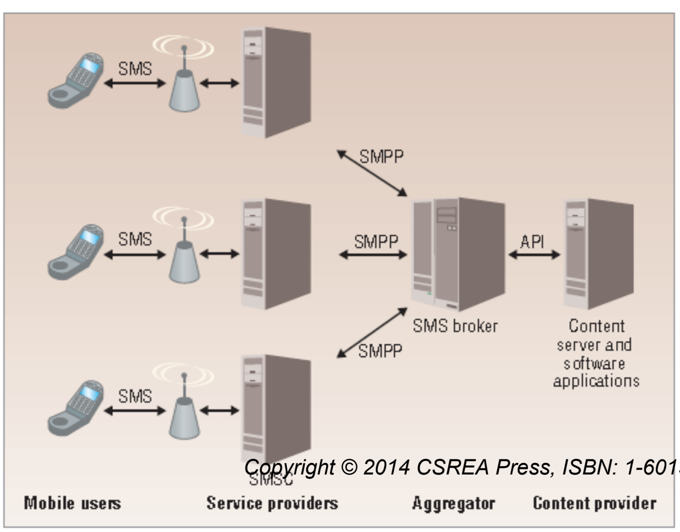
A survey of all the web applications that are available today, shows many variations. For example, the database servers can run on various platforms, as can the clients. Designers of web applications use various tools, which affect how the applications work and how they look.

In many cases, the client and server for a web application are on different operating systems. The client, for example, can be on a workstation-based operating system, such as Windows XP or Linux Ubuntu. The server for the application can also be on a workstation-based server, or it can be on an enterprise server, such as windows server. The browser uses Hypertext Transfer Protocol (HTTP) to forward user requests to a second-tier server machine. (HTTP is a communication protocol that the web uses.) The web server on the second tier invokes the local database server to satisfy the data requirements of the application.

2.3.2 Mobile phone as a tool for data collection

In the past decade, mobile phones have become an ubiquitous feature of life in developing countries. According to (World Bank, 2013), between 2005 and 2011 mobile cellular subscriptions nearly tripled in the developing world, increasing from 1.2 to 4.5 billion. In Africa, the region with the fastest mobile subscription growth rate, mobile cellular subscriptions increased from 87 million in 2005 to 433 million in 2011. The rapid growth of mobile phones can be attributed to enhanced features and user interfaces, increasing penetration of faster wireless broadband networks, increased screen size, and more attractive pricing. Mobile phones leverage the latest cellular technology and wireless Internet capabilities, providing a wide range of features such as texting messages, connecting to the Internet, use of third-party applications, e-commerce, and GPS.

One other very important feature of mobile phones which is aiding data collection is the Short Messaging Service (SMS). This basic service allows the exchange of short text messages between phones and also between phones and other applications like a web application.



Although the popularity of text messaging is well established in many countries, in Ghana, interest in the thumb-driven phenomenon has only recently skyrocketed. Though text messaging is currently limited to 160 characters, it is not dependent on making direct two-way contact with the respondent as in the case of the smartphone GPS. Thus, it can serve as a powerful survey tool for short, frequent data collection or for inviting respondents to complete surveys at their convenience (Callegaro, 2002).

FIGURE 1: SMS SYSTEM ARCHITECTURE.

The SMS system architecture depicted in Figure 1 shows that content providers go through a message aggregator rather than communicating directly with the various Short Message Service Centers (SMSCs). The message aggregator uses the Short Message Peer-to-Peer (SMPP) to maintain connections with carrier networks. An aggregator is a business entity that negotiates agreements with network providers to act as a middleman providing access to a cellular network for messaging services to third parties who have no direct relationship with the cellular network.

The message aggregator uses the SMPP to maintain connections with carrier networks. Aggregators typically provide access to their servers either through SMPP or using customized APIs written in Java, PHP, Perl, and so on. Most aggregators will also manage the rental of the common short code for clients who do not want to deal directly with the telecom company. Another important function of the aggregator is to assist with provisioning of the short code on the various carriers.

3 Methodology – System Specification and Design

The challenge in selecting and following methodology is to provide the right processes and guidance to deliver the system. In this research, the “waterfall cycle” method of the traditional system development life cycle was used.

3.1 Functional Requirements of Web and Mobile-Based Data Capture System

Requirements are defined in terms of specific behaviors or functions which require a set of inputs for required outputs or outcomes. The functional components in this system includes:

- **The Web Module**: This module handles the storage and processing of uploaded data and facilitates the viewing and analysis of the data on the web using a web browser.
- **The SMS Module**: This module is used to send and receive messages between the web module and the various stakeholders. The module needs a GSM handset to be able to receive and send the SMS messages.

These are the two most important components of the IT-Based M&E system.

3.2 Workflow Diagram

Figure 2 shows the communication flow between the web application and the project beneficiaries using SMS. In this workflow diagram, when the status of a project changes to "completed", the web application sends an SMS message to beneficiaries of the completed project using the Twilio SMS gateway. The project beneficiaries reply to the message. The replies are stored in the Twilio database. The web application picks up these replies and updates the system database.

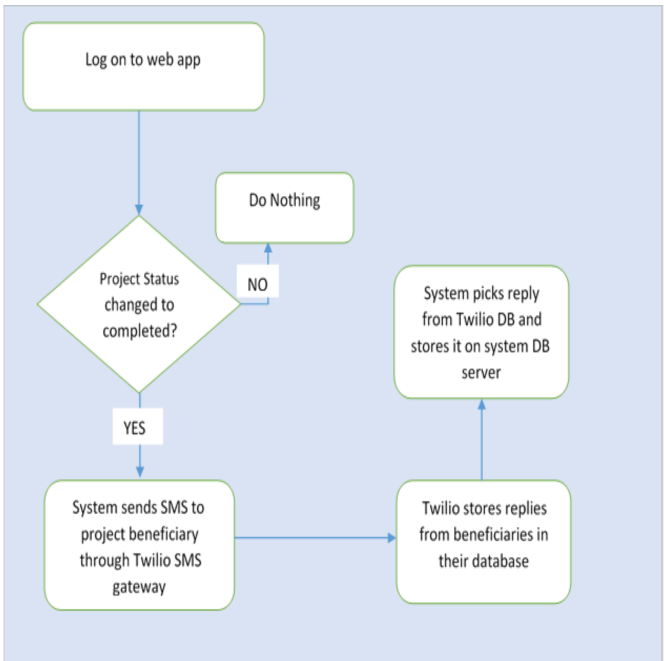


FIGURE 2 : WORKFLOW OF SEND/RECEIVE SMS BETWEEN WEB APPLICATION AND PROJECT BENEFICIARY

3.3 System Design

During the design phase, a system is designed to satisfy the requirements identified in the previous phases. The requirements identified in the requirements analysis phase are transformed into a system design document that accurately describes the design of the system and that can be used as an input to system development in the next phase. Figure 3 shows the system architecture that has been developed for this system.

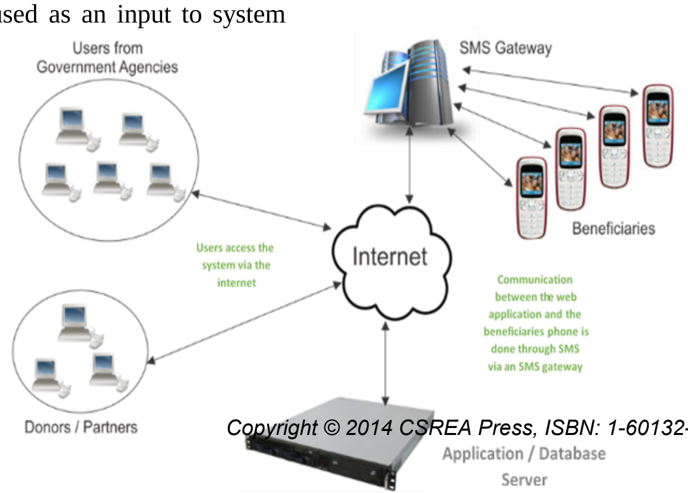


FIGURE 3: OVERVIEW OF THE TECHNICAL ARCHITECTURE

3.4 Security

Nowadays, security on the web is very essential in order to ensure data is secure and safe from intrusion during transmission and for maintaining the integrity and reliability of the data. Encryption was implemented as the key security feature. The PHP MD5 algorithm was used to encrypt data. This algorithm was mainly used during the storage and retrieval of passwords and also in providing a secure channel for updating the database. Additionally, only the administrator is able to implement changes to the system while it is offline in order for security to be effective.

4 System Development and Piloting

In this section, the various implementation techniques used are identified and discussed. The process of building and testing the various components are also described.

4.1 System Implementation

The completed system is made up of a web application which provides users with the interface to upload project data and view various project status and analysis. The system relies on the Twilio SMS gateway to provide the SMS services required by the web application to successfully communicate with the various beneficiary. There is a backend database system for permanent retrieval and storage of information. There is a real-time communication between the frontend and backend systems.

4.1.1 Development Tools

PHP: Hypertext pre-processor is a server-side scripting language that handles data submitted from the web application and SMS messages from beneficiaries. This was chosen as the server-side scripting language because of its compatibility with the Database Management System (DBMS) and interfaces with HTML to display dynamic objects.

MySQL: The DBMS used because of its low storage needs and its open source nature.

JavaScript: The primary client-side scripting language for the application, it was used for data manipulation on the client-side of the web application.

HTML and CSS: This integrated environment was used for structuring and styling the appearance of the web application.

XML: Representational language for transporting and storing data between the backend and the mapping to the frontend. It was used for transporting data between the web application and the SMS gateway.

Internet Explorer, Mozilla Firefox, and Google Chrome: These are web browsers used to test the web application.

Eclipse Integrated Development Environment (IDE): This IDE provides the tools for PHP developers creating web application. It was used in developing the web application module of the system.

Twilio REST API: This API is provided by Twilio messaging. It provided the platform for sending messages to phone numbers.

Twilio TwiML: This is an XML platform, provided by Twilio(tm). It helps in controlling messages. It is integrated into the system to provide responses to messages. It also helped to control message flow as well as trigger other application logics.

4.1.2 Web Application Login Screen

Figure 4 shows the login screen of the web application. It is the first screen displayed when the application is launched. Users are required to input their username and password to access the system.

FIGURE 4: SNAPSHOT OF LOGIN SCREEN OF WEB APPLICATION

There are four levels of users in the pilot system, and each user is assigned to a security clearance level upon creation. These four levels are explained below:

Administrator – This user has the clearance to view/edit/delete data from all Municipal and District Assemblies (MDAs) who are registered in the system. They can add users and delete users. An administrator virtually manages the frontend of the web application.

Desk Officer – Desk Officers are research officers at the Policy Unit who have the responsibility of overseeing the various categories of MDAs. They have the privilege to view/edit data from all MDAs, but are unable to perform any deletion function.

Viewer – The VIEWER is only able to view data from all MDAs. They are unable to either edit/delete records.

Respondent – Users in this category are usually stationed at the various MDAs. They are limited to the view of only the MDA they represent. They can add/edit/view records that belongs to their MDA. They are unable to view records from other MDA.

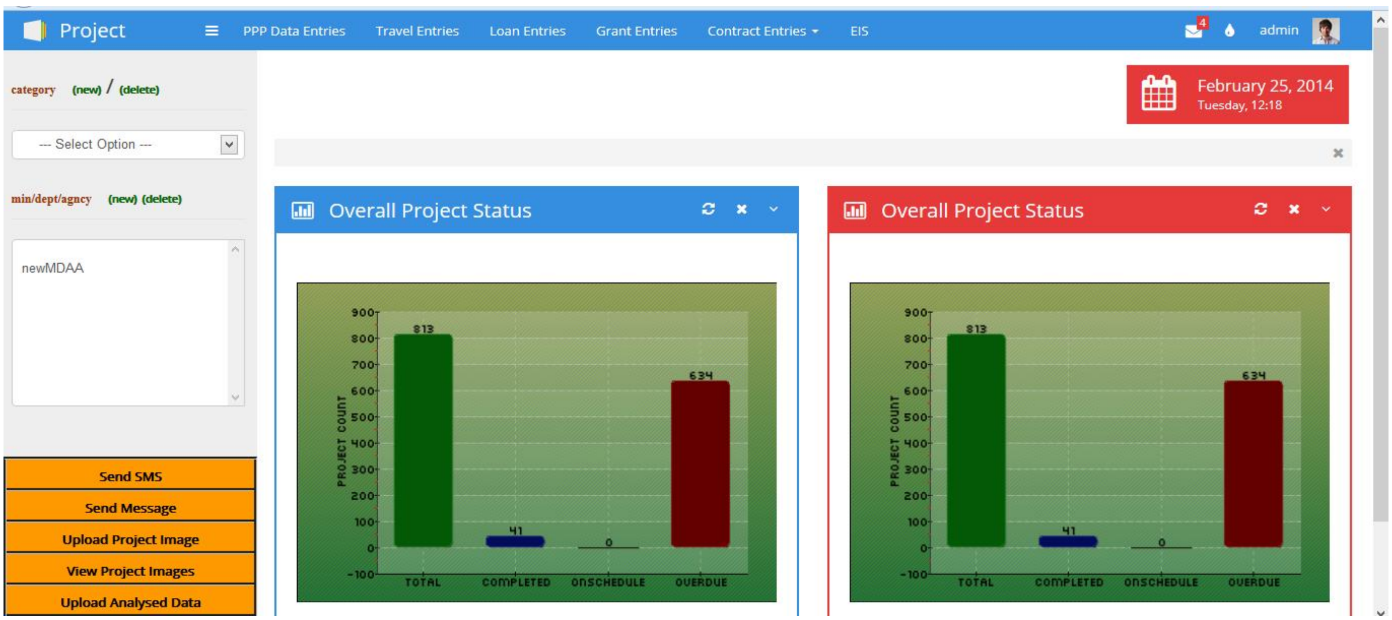


FIGURE 5 : DASHBOARD VIEW OF WEB APPLICATION

Figure 5 shows the dashboard that is seen by the user after login. It provides easy access to important functionalities such as policy, programme and project details, EIS (Executive Information System), Image Upload etc.

The EIS shows the status of a project in real-time. It depicts the status using a color code as displayed in Figure 6 below

COLOR	STATUS
blue	Completed
red	Overdue
green	On-going
yellow	At Risk

FIGURE 6 : EIS PROJECT STATUS INDICATORS

4.1.3 Twilio for Sending and Receiving SMS

Twilio SMS enables sending and receiving SMS messages programmatically between the web application and project beneficiaries through a Twilio phone number. The web application triggers an event which sends a request to the Twilio SMS gateway server when the status of a project changes to 'Completed'. The Twilio server, sends a preconfigured message to the beneficiaries of that particular project. The Twilio server receives an SMS reply from the beneficiaries and stores them in a database table. The web application then picks up the data from the Twilio database and then replicates it on the application/database server.

For this to work, Twilio exposes outbound SMS messaging functionality through its Twilio REST API as shown in Figure 7. When the web application makes a POST request, the Twilio SMS Messages resource sends an SMS message to the phone numbers specified.

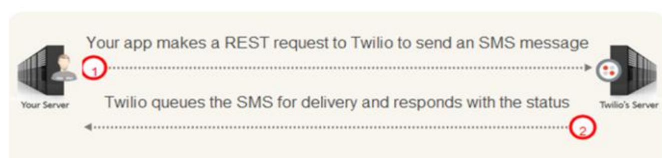


FIGURE 7: SEND AND RECEIVE SMS IMPLEMENTATION (http://www.twilio.com/docs/howto/sms- notifications-and-alerts)

4.2 Piloting

The pilot was conducted with about 300 participants from the various government ministries, departments and agencies in Ghana (as a case study). The pilot test was considered to be a success by both the project team and the policy unit. It was observed that the set of deployed technologies as well as those that were developed enabled the collection of meaningful and accurate data on the policies, programmes, projects, staff details, travels, loans, grants, and even contracts of the various MDAs at all times.

One of the main goals for this pilot was to find out whether or not the resultant data could be effectively and reliably collected using this web-based application and SMS technology. It was concluded that the system proved to be a better option in providing the Policy Unit with an improved way of performing their M&E functions

5 Conclusion and Recommendation

In conclusion, the research explored the use of ICT in a monitoring and evaluation environment. It showed through a pilot implementation that the use of ICT in M&E data collection is realistic. However implementing a web-based tracking system as well as an SMS feedback system even on a pilot scale is not an easy task, a lot more research will be required for a nation-wide implementation.

Due to the fact that the usage of the web has recently gained momentum, many of the technologies standards in this field are still in a state of change. It is therefore advisable to use open standards, or source, solution for robustness and also provide more tested technologies. It is also very important that special consideration is given to the information system design to prevent a "data avalanche" situation. Providing meaningful views to a database containing thousands of data is a challenge to the application designers not only on the performance of the application, especially in real-time mode, but also on the application logic side.

Overall, the outcomes of the project are very encouraging, but more research is required before a nationwide implementation of this kind of data collection system. This research provides a good foundation upon which to base future work.

6 Bibliography

Callegaro, M. (2002). Remarks made at Roundtable on "Mobile Phones in Telephone Surveys". Annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.

Davidson, E. J. (2009) "Improving Evaluation Questions and Answers: Getting Actionable Answers for Real-World Decision Makers." Demonstration session at the 2009 Conference of the American Evaluation Association, Orlando, Fla., Nov. 2009.

Kaplowitz, M. D., Hadlock, T. D., Levine, R. (2004). A Comparison of Web and Mail Survey Response Rates. Public Opinion Quarterly, 68(1), 94–101.

Kempf, A.M., and Remington, P.L. (2007). New Challenges for Telephone Surveys. Annual Review of Public Health, 28: 113-126.

Raftree (2013b), Using Technology for Development Project Monitoring & Evaluation: An Interview with Linda Raftree, 02-07-2013.

United Nations Development Programme (2013). Innovations in Monitoring and Evaluating Results. New York.

World Bank (2004). Monitoring and Evaluation: Some tools methods and approaches. Washington, D.C. [Online] Available at: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2013/01/17/000158349_20130117102554/Rendered/INDEX/wps6321.txt. [assessed on November 17th, 2013].

World Bank (2013). Challenges and Opportunities of Mobile Phone-Based Data Collection [Online] Available at: http://www.worldbank.org/oed/ecd/.[assessed on November 12th, 2013].

Zanamwe, N., Okunoye, A. (2013). Role of ICTs in Climate Change. International Conference on ICT in Africa

Construction and Verification of SNS Which Used the Japanese Historical Government System

Hideaki Hashimoto , Takayuki Fujimoto
Graduate School of Engineering, Toyo University
Kujirai2100, Kawagoe-City, Saitama, Japan
s36d01210025@toyo.jp, me@fujimotokyo.com

Abstract – *In recent years, many SNS exist. Especially Facebook is very popular. However, the number of friends increases too much and the time line is disorderly. Therefore, the friend who has not taken communication at all exists. Moreover, time to use SNS is increasing. Therefore, We construction and verification new SNS. This system restricts a friend's number and makes the time line smart. The purpose of this system is to shorten utility time on SNS. Information is summarized for time reduce. This system does not increase a friend.*

Keywords: SNS, Information, Gathering, Web service, Communication;

1 Introduction

A. Outline

In recent years, many SNS exist. The mainstream is Facebook and a Twitter. A social network can be made now on the Internet by SNS. However, the number of friends increases too much and the time line is disorderly. Therefore, there are the friends who has not taken communication at all. Moreover, time to use SNS is increasing and what was called "SNS tiredness" is problem. Therefore, we proposed new SNS. This system restricts a friend's number and makes the time line smartly.

The purpose of this system is to shorten utility time on SNS. Information is summarized for time reduce. This system does not increase friends. We researched the appropriate number which restricts a friend and a group. As a result, it turned out that it is about 150 people that we could take communication to some extent. This number is related to the size of man's cerebral neocortex. This number is called the Dunbar's Number. Furthermore, We referred to the traditional five-man-groups system of Japan. Hideyoshi Toyotomi made this system in Kyoto in 1597. He unified the political factions of Japan. From these research, we decided to make ten five-man groups. This number of friends let's you have deep communication. This paper is about the development of gathering friends.

B. Purpose

The purpose of this paper is a reduction in utility time of SNS. This paper proposed new SNS. It differs from the existing SNS. Number of friends are restricted and grouped. Moreover, this system summarizes information by the friend. This SNS supports communication in everyday life. It does not make a new friend. We perform construction and verification of this system.

2 Background

C. Social Networking Service (SNS)

A social networking service called SNS is an online service. The purpose is promotion of communication. SNS simplifies communication steps with people who have common hobby. Moreover, it supports to have communication with a friend. In recent years, the user can participate freely. SNS are: Message trading, photograph contribution, and a profile. Many SNS are for free. A business model is advertising revenue. SNS was born in the United States of America in around 2003. Facebook and Twitter are the most popular.

a. Facebook

Facebook is the world's largest SNS. Users are about 800 million people. Facebook was developed for students in the United States of America. The difference from other SNS is real name registration system. A user generally exhibits a mug shot. Facebook supports communication with friends. Facebook has a "Like Button". Which button simplifies communication between users.

b. Twitter

Twitter is micro blogging service. The users send the messages up to 140 words. As for Twitter, English version was started in 2006. Twitter showed up Japan in 2008. This service can make friends anyone. Moreover, this service can acquire new information. Users are 500 million or more today.

c. LINE

LINE is a proprietary instant messaging application for smartphones. As of 2012, LINE has more than 49 million users in over 230 countries. LINE can communicate for free.

Moreover, Have a chat function. A chat function has a sticker, which shows you a feeling. LINE import the address. It is supports communication in everyday life. In recent years, LINE spread quickly. The purpose resembles this paper.

D. Research

a. Japanese traditional five-man groups system

GONIN-GUMI is Japanese traditional five-man groups system. It is Command and Control system of Japan. Hideyoshi Toyotomi made this system in Kyoto in 1597. He unified the political factions of Japan. This system enabled private sector control. The representative was called GONIN-GUMI-GASHIRA. This is a famous historical system of Japan.

b. Dunbar's Number

Dunbar is the British anthropologist who proposed Dunbar's Number. It is a suggested cognitive limit to the number of people with whom one can maintain stable social relationships. The number is 150 people. It is related to the size of a cerebrum.

Moreover, the degree of intimacy can be expressed with a concentric circle. 5best friends go in to the circle nearest to the center. They are un removable following 10 people, next to the best friend group go into the outside of center group. Following 30 people and 100 people group, which separated by how close to you go out circle of circle in order. As group go to outside of the center, this system will arrange us how often it try to contact with our friends, once a week, once a month, once a year and so on.

Namely, these are reflecting the degree of intimacy. When the newest person in a central circle enters, exchange is performed.

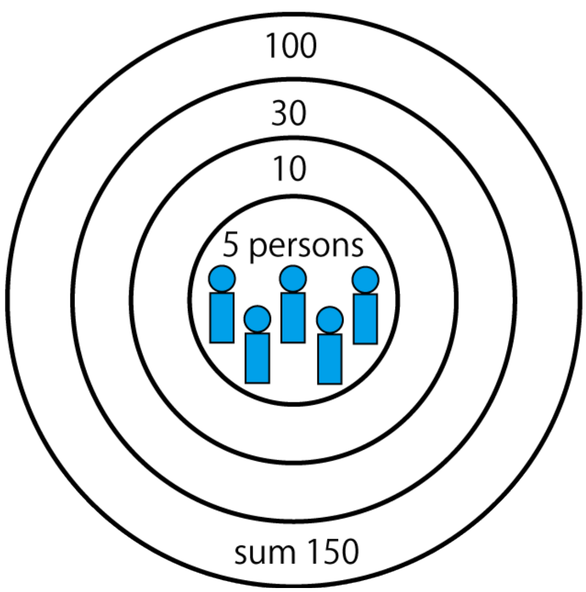


Figure 1. The concentric circle of the degree of intimacy

c. Group of everyday life research

There are many groups in everyday life and it is thought that the number with man could take communication at once has a limitation. For example, a family restaurant has a box seat for 4~6 people. This number is good for taking. Around people circle is suitable for Buffet-style party. The conversation became difficult, when there are so many people in a group.

Moreover, 5~9 people are suitable for Agile development. This number can be developed as functionally as a sport team. To others, business organizational theory must not exceed 150 people. This is because sick leave increases in case of a lot of people. In the case of 150 or more people, a grading system is required.

At the last, in the case of an army, number of the minimum units are 5 people.

3 MECHANISM OF A GATHERING SNS SYSTEM

E. Abstract of a system

This system can register up to total of 50 friends.
(5 people * 10 groups = 50 people)

A user takes communication by substitute timeline of a group. The member registered into the group can use substitute timeline. Others cannot use, so can keep the privacy. This system has gathering button. It can be displayed on main timeline of registration member. Namely, the information selected carefully is displayed on a user's timeline. Moreover, this system has a profile and a direct message function.

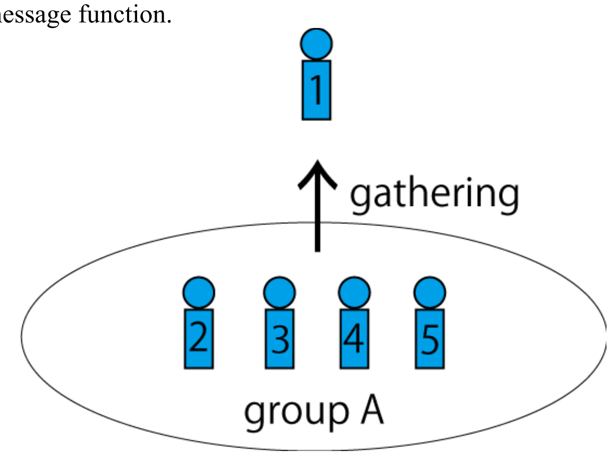


Figure 2. Image of gathering

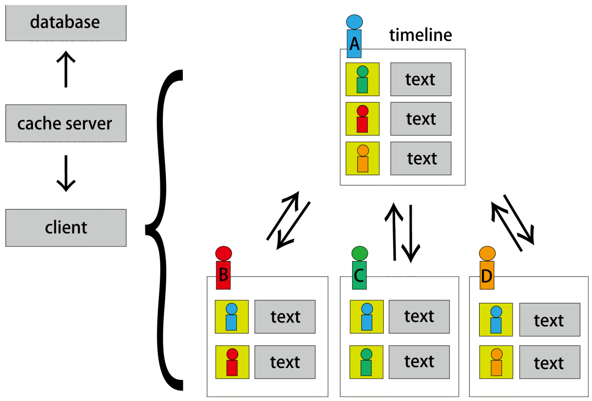


Figure 3. Structure of updating

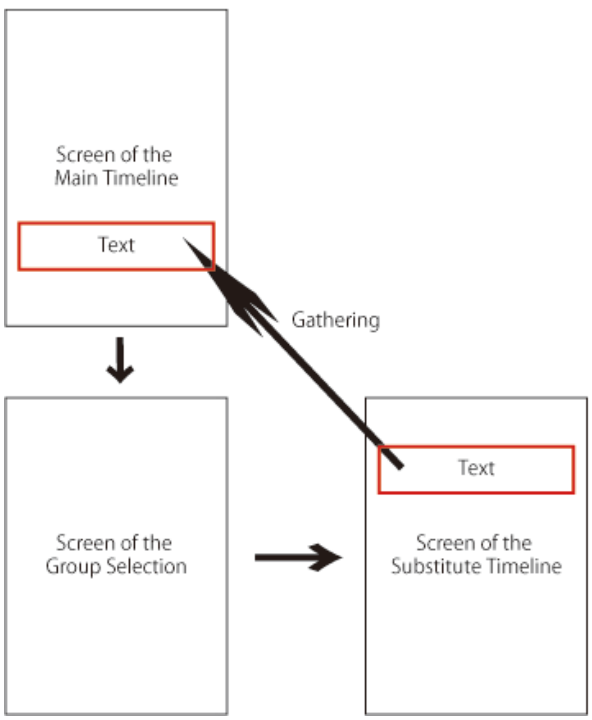


Figure 4. Step of operation

F. *Difference from the existing SNS*

A big difference is the gathering button. This paper supports communication with a friend, and restricts the registration number. Gathering of the information is carried out. The existing SNS enabled borderless communication. This system makes a new friend. However, the user could not process information.

TABLE 1. Function, Usefulness, Comparison list

	Facebook	Twitter	LINE	This paper
Name	Real Name	Anonymity	Real Name	Real Name
Friend	Check	No Check	No Check	Check
Privacy	Setup is possible	Setup is possible	○	○
Response	○	◎	○	○
Quality of information	△	×	○	◎
Utility time	△	×	○	◎

G. *Validity of 5-man-groups mechanism*

Gathering SNS system proposed by this paper has validity. It was researched and proved. Information gathering is easy for 5-man-groups. This system supports communicating with a friend. Moreover, it can keep the privacy. People have about five communities. Namely, the number 5 is the optimal.

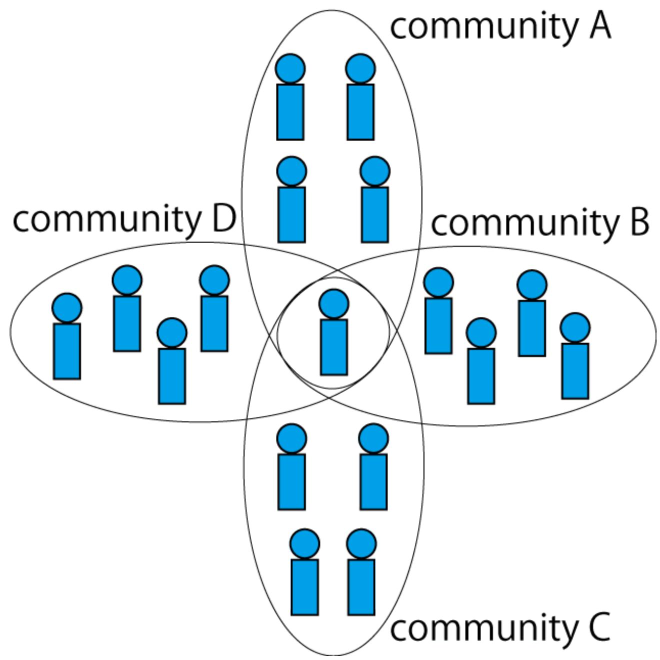


Figure 5. People's community

H. *Details of a mechanism*

a. Various screens

First, this system was developed using HTML, PHP, and MySQL. SNS today enabled to borderless communication.

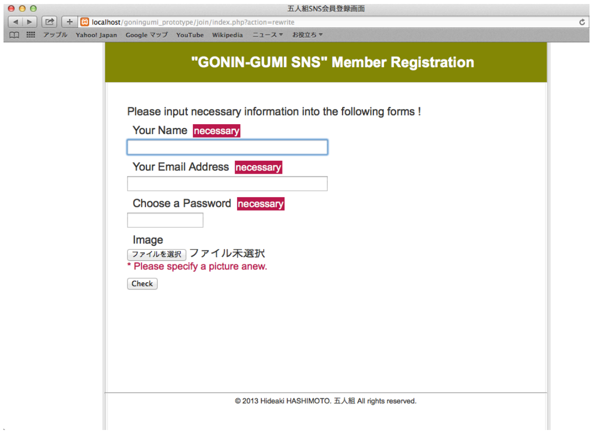


Figure 6. Screen of the member registration

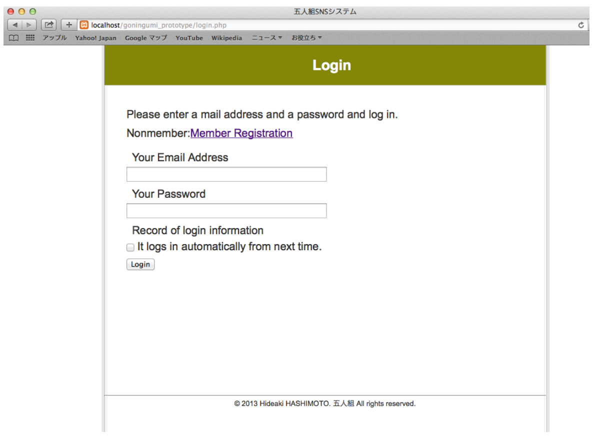


Figure 7. Screen of the login

Main timeline is home screen. Important information will display on here. Information is acquired even if a user does not go into the room of a group. Gathering SNS system can be used in short time. The image figure was attached downward. "GONIN-GUMI-GASHIRA (main timeline)" is displayed on the upper part of a screen. The central part is main timeline. If the report is clicked, it will move to substitute timeline of the group to which a contributor belongs. A contributor's icon is displayed on left side. "GONIN-GUMI-GASHIRA" is home button. If button is clicked a screen will change.

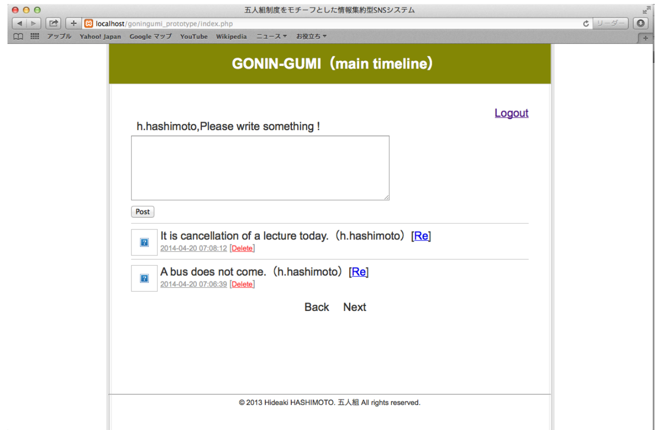


Figure 8. Screen of the main timeline

The figure attached downward is a group selection screen. It has 10 groups (Group 1 ~ Group 10). A user chooses the room of a group here. A group is named and it is made intelligible. " GONIN-GUMI-Group" which shows group selection screen is displayed on the upper part of a screen. The central part is the group selection column. This will move to the room of a group, if a group name is clicked.

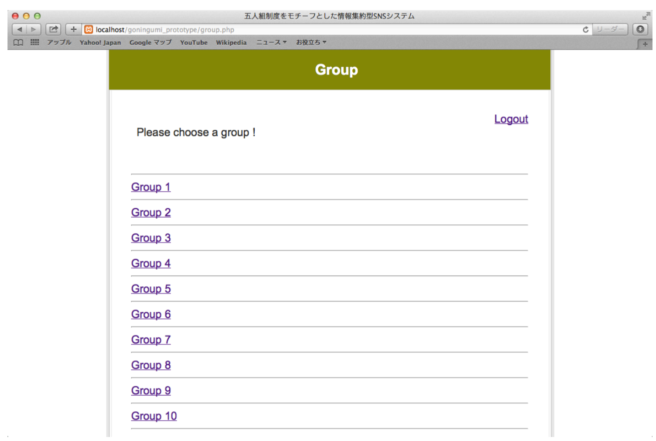


Figure 9. Screen of Group selection screen

Finally, the figure attached downward is substitute timeline. This is a room of the group and only for the member. Here, it communicates by 5 people. A group name is displayed on the upper part of a screen. The central part is the substitute timeline. gathering button is on the right-hand side of delete button. It is very important.

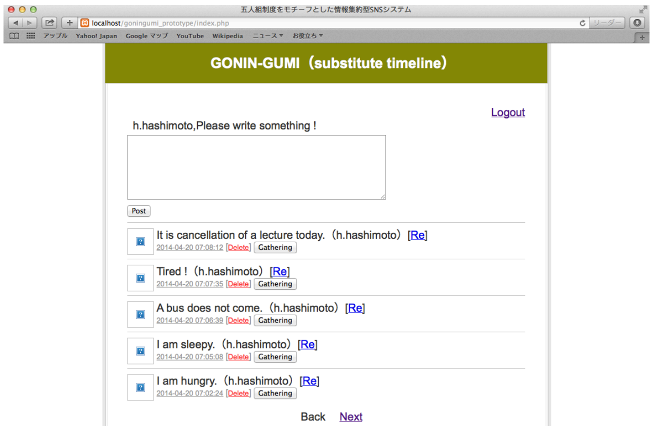


Figure 10. Screen of the substitute timeline

b. *Gathering method*

The method of the gathering is user dependence model. Gathering button is pushed to important information. This is displayed on a member's main timeline.

Validation methodology

Scale: 5 people

Subject: Please decide on the place and day of the week of a meal meeting using this system.

Questionnaire

① Has it determined early?

② Is there little amount of information of the main time line?

③ Is the quality of the information displayed on the main timeline good?

④ How many times did you use the gathering button?

⑤ Did you feel "SNS tiredness"?

The result of an easy subject experiment

TABLE 2. ①Has it determined early?

very late	late	Usually	early	very early
0	0	1 people	4 people	0 people

TABLE 3. ② Is there little amount of information of the main time line?

very large	large	Usually	small	very small
0	0	1 people	2 people	2 people

TABLE 4. ③ Is the quality of the information displayed on the main timeline good?

very bad	bad	Usually	good	very good
0	0	0	2 people	3 people

TABLE 5. ④ How many times did you use the gathering button?

not used	1 time	2 times	3 times	more
3 people	1 people	1 people	0	0

TABLE 6. ⑤ Did you feel "SNS tiredness"?

felt very much	felt much	Usually	did not feel	did not feel at all
0	0	0	4 people	1 people

Verification result

It judges from TABLE 2.,
It can be said that this system can determine a schedule early in general compared with the existing SNS.

It judges from TABLE 3.,
It can be said that there is in general little amount of information of the main timeline.

It judges from TABLE 4.,
It can be said that the quality of the information on the main timeline is in general good.

It judges from TABLE 5.,
It can be said that the use frequency of the conclusion button was comparatively small, and it was appropriately used so that we might recommend. It was used only for determination matters, such as a place and time.

It judges from TABLE 6.,
It can be said that this system does not feel SNS tiredness.

4 Results

This paper restricted and grouped the registration number of SNS. Namely, Gathering SNS system has smart timeline. We developed new SNS which can be used in short time. This research is useful as Table 1 showed. 5-man groups have validity as the number of Dunbar showed them. Since this SNS has gathering button, it is new. We performed construction and verification of this system. It is thought from a verification result that this system can achieve the purpose.

5 Conclusion

This paper's theme is [Gathering SNS system by Japanese traditional 5-man-groups “GONIN-GUMI”]. This is because we thought that the increasing friends are the problem. However, having communication with friends in actual world is difficult. We thought that the improvement of SNS was required.

6 REFERENCES

[1] Robin Ian MacDonald Dunbar “How Many Friends Does One Person Need?” ‘Faber and Faber’ 2010

[2] Seiichiro Kusunoki “Edo No Ofregaki Syourui Awaremi No Rei Kara Ninsousyo Made” ‘Gurafusya’ 2008

[3] Craig Larman “Agile and Iterative Development” ‘Nikkei BPsya’ 2004

[4] Yomiuri Shimbun 2012/7/17/P11

[5] Akiko Orita, "SNS Ni Syuuyaku Suru Zyouhou : Networking Kara Lifelog (<Tokusyu> Social Networking Service Katuyou Shinan)",Zyouhou No Kagaku To Gizyutu 61(2), 70-75, 2011-02-01

[6] Akihiro Kasuga, Yuuiti Saegusa ,Younosuke Furui, Haruo Hayami,"SNS deno chat ni yoru tomodati no wa kakudai shien system no teian(Session 4: Relation with people and a person)",Jyouhousyori Society Report of research. GN, [Groupware and Network Service] 2007(32), 61-66, 2007-03-22

[7] Shigeki Toyama, “Syakou Wo Dezain Suru : SNS Wo Riyo Shita Chiiki Syakouba No Sousetu No Kokoromi” Zyouho Bunka Gaxtukai Shi 18(1), 11-18, 2011-07-31

SESSION

BIG DATA, DATA MANAGEMENT AND STORAGE + XML + PRIVACY ISSUES + APPLICATIONS

Chair(s)

TBA

Big Data Analytics Framework for Improved Decision Making

Brian H Xu[1], Sathish Alampalayam Kumar[2]
[1]Honeywell Aerospace, USA - brianxu8@gmail.com
[2]Coastal Carolina University, USA – skumar@coastal.edu

Keywords: *Our big data technologies based analytic framework blend rapid data processing capabilities of computers with intuitive decision making skills of humans. Current Information Technologies (IT) and computing clusters make pertinent or contravening information elusive to almost all decision makers. Our scalable machine learning algorithms are capable of performing autonomous data discovery and analysis operations across various data sources of structured and unstructured data types. Specifically, based on this framework, human decision making workload would be significantly reduced. Our philosophy is that human-machine interactions are to be optimized by autonomously gathering relevant information for decision-makers and drawing meaningful real time conclusions from the structured and unstructured big data. The experimental results demonstrate that our Big Data analytics framework based Intelligent Agents are more competitive with the baseline algorithms and produce better rule induction algorithms with higher predictive accuracy.*

Keywords: Intuitive Decision Engine, Big Data Analytics, Cloud Computing, Intelligent Agents, Machine Learning.

1. INTRODUCTION

A Computer can process vast amounts of data in a short time but programming them to detect patterns is extremely difficult. Humans, on the other hand, are adept at extracting patterns from large amounts of data, yet are limited by the rate at which they can apply these cognitive abilities to large amounts of quickly changing information [15]. Consequently, computers have low success rates when tasked with detecting relevant and important information embedded in complex data streams, while humans are unable to effectively process large amounts of data that are fast becoming a hallmark of today's decision-making environments. Compounding the problem, traditional support technologies, like decision aids and visualization tools, create an artificial barrier which plays to either the strengths of computers or the strengths of their human operators, but not both. This problem could be addressed by imparting human intuitive skills through intelligent agents to the computers. Intuition is the process through which humans quickly make sense of partial, incomplete, or rapidly presented information. Intuitive decision-making processes are associated with rapid recognition of patterns among incoming streams of information followed by retrieval of associated knowledge without conscious attention [17]. The neural processes underlying intuition occur on a very rapid timescale and can be simulated through intelligent agent technologies.

In this paper we present the big data analytics framework for improved decision making to augment the current data processing techniques. This will enable humans to rapidly analyze large amounts of information in complex information environments. This framework is based on the principle that higher the amount of information processed per unit time, lesser is the overall decision-making time.

In Section 2, we have discussed the related work that is ongoing in this area. In Section 3, we have explained our Big Data Analytics Framework and Technologies. Section 4 discusses our prototype implementation and preliminary experimental results. In Section 5, we present our conclusion and future direction.

2. RELATED WORK

In recent years, a number of data mining algorithms and knowledge discovery technologies have been developed and applied for finding information from the ever-growing various data sources such as structured, unstructured, relational, temporal-geospatial datasets, etc., Hadoop [1] and Cloud Map Reduce [2, 8, 16] have been the leading parallel and distributed computing technologies that are used in wide variety of applications and domains. However, these existing technologies lack with respect to autonomic intelligent information processing operations across multiple modalities of the data [13, 14].

State-of-the-art MapReduce computing algorithm is implemented on parallel and distributed real time Hadoop or Storm [12] computing platform, which is used to support search engines, recommendation engines and big data analytics in wide variety of applications. However, Hadoop-MapReduce technologies alone are inadequate to perform autonomous data discovery and analysis operations across various data sources (For example, Department of Defense Global Information Grid, Human Intelligence Collection, Open Source Intelligence Collection etc.) and data types (e.g. text, video, audio, images, etc.). This is due to the reason that they are incapable of seeking, understanding and presenting information and moreover both Map and Reduce functions are implemented on fixed nodes.

Cloud Hadoop is a framework built for capturing, organizing, storing, and analyzing data from multiple sources. A growing number of commercial and military organizations using Cloud Hadoop have found it to be an indispensable tool to improve decision-making in real time and gain a competitive edge in decision making. But many users have encountered its serious limitations to process data in real time. Classic Hadoop also needs novel machine learning algorithms to meet the needs [15, 17].

To address these limitations, our objective is to create a real time Big Data Analytics framework by synergistically integrating our innovative Intelligent Agents with real time Hadoop framework, streaming data processing tools, and proven Machine Learning (ML) algorithms. We have utilized the open source Mahout, Rapid Miner, and Weka for the machine learning algorithms, and leveraged our intelligent agents and Blackboard pattern enabled agents [7, 9]. The proposed framework has the following benefits:

1. Automated and scalable real time data-to-decisions capability by developing and integrating the networked and distributed mobile Agents, scalable cloud technologies, Cloud MapReduce, machine learning algorithms and Storm (real time data stream processing).
2. Autonomous decisions and analysis of large data sets by using cloud parallel computing and novel machine learning technologies such as Naive Bayes, Support

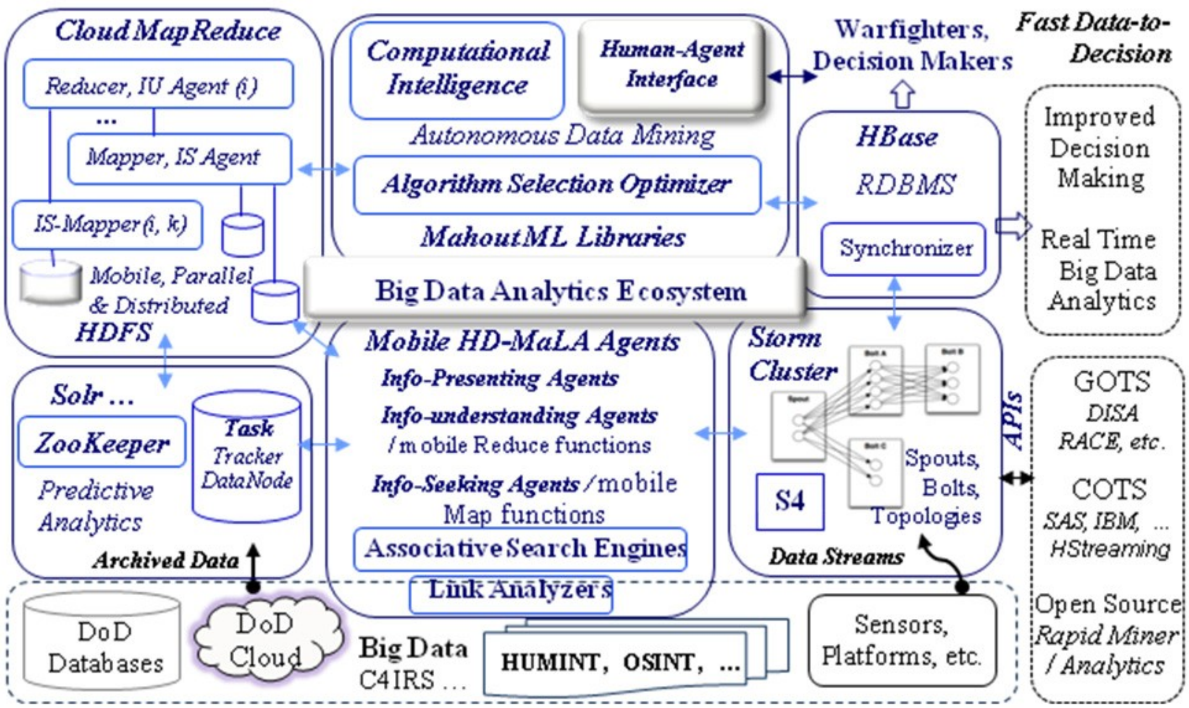


Figure 1. HD-MaLA Big Data Analytics Framework

3. Vector Machines (SVM), and Principal Component Analysis (PCA).
4. Scalable computing algorithms using intelligent mobile agents with machine learning algorithms to needs for autonomous system operations, make smarter decisions and draw keener insight and meaningful conclusions from massive and heterogeneous datasets.
5. Automated data mining and knowledge discovery performed by intelligent mobile Agents that are closest to the task related data sources, in order to provide distributed data mining and autonomous Data-to-Decisions with minimal computing and networking resources (without moving huge data to analytic codes).

3. BIG DATA ANALYTICS FRAMEWORK

To address the existing limitations discussed in section 2, we have designed and integrated real time Hadoop-MapReduce frameworks with distributed machine Learning enabled intelligent Agents (HD-MaLA). The proposed big data analytics framework integrates next-gen Hadoop with our intelligent mobile agents and machine learning algorithms on top of the upper layers of the big data stack: Hadoop Distributed File System (HDFS), HBase, Apache Solr and ZooKeeper. Since cloud computing technologies have been successfully demonstrated for real time applications, they will be leveraged to develop our framework. Specifically, various problems are easily

As shown in Figure 1, the novel HD-MaLA architecture is designed by creating and integrating cloud computing technologies, intelligent mobile agents, machine learning algorithms, and algorithm selection optimizer. HD-MaLA solutions have key benefits over current approaches due to the following reasons: (1) It deploys novel Intelligent Mobile Agents powered by machine learning algorithms and technologies; (2) Integrates and leverages existing off the shelf tools and products via Application Programming Interfaces (APIs), and (3) Implements parallel and distributed complex analyses using cloud MapReduce paradigm. As shown in Figure 1, major HD-MaLA components and technologies are integrated to significantly increase synergistic capabilities for real time big data analytics. Technical details of major HD-MaLA components are briefly described in the following.

Real time Hadoop (Storm): Hadoop is an open source, flexible Java framework for large-scale data processing on commodity hardware networks and it has become the industry de facto framework for large data processing. To make full use of computer resources, it is important to optimize performance, including CPU, memory, and I/O (both disk and network). For real time applications, Real time Hadoop implemented by Apache Storm will be used to develop HD-MaLA architecture. Cloud MapReduce, machine-learning algorithms, and HD-MaLA Intelligent Mobile Agents are supported by Real time Hadoop and Storm clusters.

HDFS and HBase: HDFS is designed for storing and sharing files across wide area networks. It runs on commodity hardware and provides fault tolerance, resource management, and most importantly, high throughput access to application data. HBase provides large quantities of data in the order of billions of rows and millions of columns. HBase is essentially a key-value store with efficient indexing on key access, a semi-structured data model for value representation, and range-search capabilities supported by key ordering. HBase on top of HDFS helps the HD-MaLA to manage big data (volume, velocity, and variety data). The HD-MaLA system will be developed on top of HDFS and HBase to store and manage the data in our Hadoop cluster. Using HDFS, over 10, 30 and 100 percent performance improvement are expected for the sort, random

expressible as MapReduce computations. In fact, MapReduce has been successfully used for the generation of big data in NoSQL databases such as HBase, MongoDB etc., for web search services applied in sorting, data mining, machine learning, and social networks. The addition of MapReduce to the regular SQL interface enables fast, large-scale analytics over Big Tables such as Greenplum, HBase, etc., to reduce query times by several orders of magnitude.
write and sequential write benchmarks respectively on commodity computers in our Hadoop clusters [1, 5].

Cloud MapReduce: A large variety of problems are easily expressible as MapReduce computations. Specifically, Cloud MapReduce has been successfully used for the generation of data for production Web search services [2]. Cloud MapReduce will be integrated into HD-MaLA architecture to significantly accelerate machine learning and information seeking and enable fast large-scale analytics, autonomous data-to-decisions, and distributed data-mining.

Machine Learning Algorithms: Currently, non-cloud machine learning libraries lack scalability, are not well tested and not built with production-quality, and they also lack open source community support. Therefore, Mahout has been created with scalable machine-learning algorithms all written using Map-Reduce paradigm, including (a) supervised learning algorithms such as neural networks, support vector machines, Naive Bayesian classifiers, decision trees, random forests, and logistic regression; (b) unsupervised learning algorithms including k-means, hierarchical clustering, self-organizing maps, fuzzy k-means, Dirichlet distribution, PCA, Independent Component Analysis (ICA), Expectation-Maximization, and Mean-Shift. All these ML algorithms will be integrated in our HD-MaLA tools and used by Info-Seeking and understanding Agents to mine data and discover actionable information. Hundreds of ML algorithms have been tested by Mahout libraries, Weka, and Rapid Miner. Best ML algorithms are automatically selected by our Algorithm Selection Optimizer to address specific needs of the HD-MaLA users. The selected machine learning algorithms are executed by Intelligent Mobile Agents in Map-Reduce paradigm to enable computational intelligence, real time big data analytics, and scalable computing algorithms.

HD-MaLA Algorithm Selection Optimizer: The HD-MaLA Optimizer consists of Genetic Programming (reproduction, crossover and mutation.), Grammar, the individual representation, the population initialization process, and the individual evaluation procedure. The effectiveness of the HD-MaLA Optimizer has been successfully evaluated in automatically designing rule induction algorithms, when the Optimizer is fed with data across different application domains. Proper machine learning algorithms are selected by the Algorithm Selection

Optimizer to solve specific data mining problems such as, classification, clustering and regression.

ZooKeeper and HD-MaLA Classification Server: To support Mahout-based services in big data analytics, ZooKeeper is used in HD-MaLA tools. ZooKeeper allows us to connect to a small cluster of servers that provide an API to access what looks a lot like an ordinary file system. In addition to very simple create, replace, and delete functions, ZooKeeper provides change notifications and a variety of consistency guarantees, that simplify building reliable distributed systems even in the presence of server failures, maintenance periods, and network partitions [12]. ZooKeeper and Mahout are used for classification in HD-MaLA as shown in Figure 2.

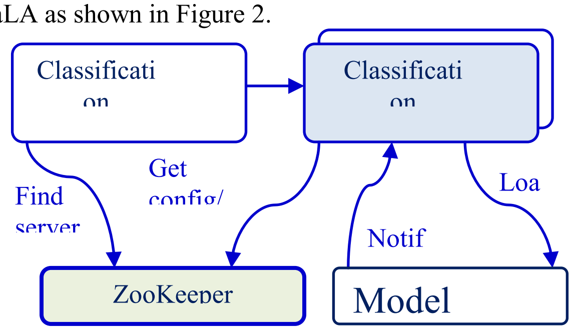


Figure 2. Mahout Classification Supported by ZooKeeper in HD-MaLA.

The arrangement shown in Figure 2 informs the ML classification server of where to find the classification model definition and allows the client (human and software agent) to find all active servers. A server queries ZooKeeper to find out what model to load. After loading that model, the server is ready for traffic by writing a file into ZooKeeper. Whenever the file indicating which file to load is changed, Zoo-Keeper alerts all servers of the change. On the other hand, when a client queries a server, it can look into ZooKeeper to find out which servers are currently providing service and pick one server at random. The overall savings due to these changes are about 3 to 4 orders of magnitude and they provide a powerful extension of the way Mahout can be used for classification [12]. More specifically, when serializing stochastic gradient descent ML models for deployment as classifiers, it is advisable to only serialize the best performing submodel from an Adaptive Logistic Regression. The resulting serialized file will be 100 times smaller than would result from serializing the entire ensemble of models contained in the Adaptive Logistic Regression object.

Intelligent Mobile Agents: The Intelligent Mobile Agents in our HD-MaLA framework perform specific tasks in the following ways:

(1) Info-Seeking Agents proactively move to data sources cross networks (of computer clusters, clouds, etc.) based on the paths selected and optimized by our Algorithm Selection Optimizer in the Mahout libraries. Info-Seeking Agents has the following features. a) They are designed and developed by leveraging source codes (in Java) of our real-time intelligent agents and our Blackboard pattern enabled agents [6, 7]; b) They move to task-related data sources (e.g., datasets from sensors and platforms, etc.) and work as mobile Map functions; and c) They send the processed data to the Info-Understanding Agents that use the reduce functions to generate the discovered information for the warfighters.

(2) Info-Understanding Agents automatically mine and discover actionable information from various data sources by using our parallel and distributed machine learning algorithms that are optimally selected by HD-MaLA Algorithm Selection Optimizer from open source ML libraries such as Mahout, Rapid Miner, SAS, R ,Weka etc., and emerging algorithms created for specific applications. Info-Understanding Agents are developed based on C++ source codes of our integrated Link Analyzers and Associative Search Engines [11].

(3) Info-Presenting Agents intuitively present the sought information to specific users based on their experiences, skills, specific tasks, and adaptively display the results according to the users' devices (e.g., tablet computers, mobile phones, etc.).

For predictive analysis, these HD-MaLA Agents conduct real time Map-Reduce jobs, look and discover information and understand information. Technically, Map and Reduce functions are dynamically created and implemented by MaLA Info-Seeking Agents and Info-understanding Agents across networks (computer clusters, clouds, etc.), based on users' specific missions or tasks. HD-MaLA mobile Agents mine data and discover information with no or little data movement, minimized time for disk seeking and I/O accessing, by minimizing data movement while supporting both batch and online real time applications.

HD-MaLA Agents understand the information by leveraging our proven Link Analyzers and Associative Search Engines. As known, operational adaptability is required at the operational level of real time decision making scenario, which requires cognitive dominance and decentralized decision making. HD-MaLA agents provide such decentralized decision making capabilities and enable real time operational adaptability.

In general, HD-MaLA Intelligent Mobile Agents

a) Distribute mobile Map and Reduce jobs using Hadoop configuration interface

b) Dynamically adjust the MaLA mobile Agent based Map Reduce analytics and functions using the real time Hadoop / Storm configuration interface based on new hints and updated system cost and loads
c) Retrieve intermediate MapReduce execution results and summarized execution status
d) Perform lifecycle management of MapReduce Agent assignment

As shown in Figure 1, HD-MaLA tools seek, understand and present information from both archives databases (HBase, SQL, MySQL, etc.) and real time data streams such as military sensors and platforms for the following functions:

(1) To process streaming data, input data streams are continuously mapped on arrival and the reduce step operates on moving time windows of data. Streaming data processing by S4 in HD-MaLA supports the complex event processing functions typically found in data processing platforms for data-to-decision, including filter, sort, correlation, aggregation, pattern matching, projection and user-defined functions.

(2) To process archived data, Info-Understanding agents move across data clusters and clouds to the data source to execute the Reduce functions for summarizing the Map function outputs from Info-Seeking Agents. These HD-MaLA Agents use the least computing resource consumptions (CPU, memory, bandwidth, etc.). For instance, HD-MaLA Agents (about 1 KB – 2 KB per agent) move to the big data sources (e.g., 100 Terabytes per source), process the data locally and send back the Map function outputs (e.g., 0.2 – 0.5 KB) individually to the Reduce functions.

Info-Understanding Agents perform advanced analytics by using machine learning algorithms selected by the Algorithm Selection Optimizer as mobile Reducers for fusing information from Info-Seeking Agents as mobile Mappers. Both the Map-Reduce functions of the HD-MaLA Agents are implemented in Hadoop paradigm. Our mobile parallel analytics is implemented by HD-MaLA MapReduce Agents that can be applied to various data sources across networks and clouds in order to enable real time big data analytics. Key advantages of the Agents include least computing resource usage, improved productivity and quality, reduced costs, and increased profits.

Two key innovations of our mobile MapReduce parallel analytics capabilities are based on mobile HD-MaLA Agents to (a) move advanced analytics to the related data stores rather than moving the huge data to the central computing center; (b) process and analyze streaming datasets, sourced from sensors and platforms in real time based on open source tools.

Our HD-MaLA Agents have basic autonomous capabilities to (1) function with little or no user intervention; (2) communicate among themselves and their users, and other agents; (3) find and move to various data sources across a cloud or a computer network; and (4) react to their environments or events without instructions from their users. They also have key advanced capabilities to (5) seek information by finding (where and what) data sources directly related to the tasks of the agent users; (6) understand information by mining data, discovering relationships among concepts, recognizing patterns, analyzing and understanding topics (from mission tasks); and (7) presenting information in an intuitive way based on users background (e.g., experiences, skills, specific tasks, etc.), by using new coding techniques such as AJAX, HTML 5 etc., and by displaying the results adapted to the users' devices such as tablet computers, mobile phones, and other devices.

HD-MaLA Agents is developed using the source code from our intelligent agents and Blackboard pattern enabled agents [6, 7] to automate tasks of seeking, understanding and presenting information. Their machine learning capabilities is enabled by using open source Rapid Miner and R for desktop applications and by using open source Mahout and Weka for scalable and cloud computing environment. Mobility of HD-MaLA Agents can be implemented as mobile Mappers and Reducers by using unanchored and/or mobile objects in C++, C# or Java socket programming and object serialization. Thus, machine learning algorithms empowered by mobile MapReduce functions are implemented in Hadoop clusters to enable advanced computational intelligence and scalable computing algorithms.

4. EXPERIMENTATION AND RESULTS

4.1 Prototype of HD-MaLA Big Data Analytics Framework

In our HD-MaLA big data analytics framework prototype, we applied cloud Hadoop framework for capturing, organizing, storing, and analyzing data from multiple sources. To effectively process real time streaming data sets, Storm is integrated in the HD-MaLA system as illustrated in Figure 1. Storm is an open source distributed real-time computation system [4]. Storm makes it easy to reliably process unbounded streams of real time data processing, similar to what Hadoop does for batch data processing [3]. Storm has many use cases for testing in the following: real-time analytics, online machine learning, continuous computation, distributed remote procedure calls and Extraction Transformation Loading (ETL) etc., Storm is extremely fast with over a million Tuples processed per second per node in a huge Hadoop cluster. It is scalable, fault-tolerant, guarantees that input data can be processed, and it is easy to set up and operate. In both military and commercial applications, many data sources including Open Source Intelligence (OSINT) and Human Intelligence (HUMINT) data archives are used that includes stream datasets. Using the HD-MaLA Storm as shown in Figure 1, data streams consist of many Tuples (e.g., "search123", "target1@fields3.com"); and Spouts read from Kestrel queue and streaming APIs, Web Logs, API calls, Event data; Bolts process input streams and create new streams, they provide functions, filters, aggregation, joins and database access; Topologies are networks or directed graphs of Spouts and Bolts.

For deployment of the HD-MaLA tools in huge systems, the following four steps are needed: (1) Scope out the problem, (2) Optimize feature extraction as needed; (3) Optimize vector extraction as needed; and (4) Deploy the scalable classifier service. For tens of millions of datasets, it becomes more and more important to use scalable MapReduce based HD-MaLA technologies in order to effectively use Mahout ML algorithms (e.g., feature extraction, classifications, etc.).

Based on our prototype, Storm cluster can be developed in the HD-MaLA system, which is composed of a master node and worker nodes. The master node runs a daemon Nimbus which is responsible for distributing code, assigning tasks, and checking for failures. Each worker node runs a daemon Supervisor which listens for work and starts/stops worker processes. Nimbus and Supervisor daemons are fail-fast and stateless, which makes them robust, and coordination between them is handled by Apache ZooKeeper in the HD-MaLA cloud.

We have developed a novel cloud infrastructure for big data analytics. To ensure successful development of the proposed HD-MaLA approach, C++ source code of our Associative Search engines is reused to develop HD-MaLA Info-Seeking Agents, while Java source code of our Link Analyzers is reused to develop HD-MaLA Info-Understanding Agent, respectively [10, 11]. We will leverage and improve this cloud using Mahout ML libraries and HBase to make the cloud based HD-MaLA Agents smarter and faster. The reason why the HD-MaLA will use the HBase is that it is a scalable, distributed database and supports real-time access large relational data repositories. Our framework will need random, real-time read and write access to huge data HBase tables can work with relational databases and achieve the highest speed in processing and analyzing the big data. HBase will be built in the HD-MaLA with the following technical components:

1. Convenient base classes for backing Cloud Hadoop MapReduce jobs and functions with HBase tables
2. Query predicate pushed down via server side that scan and get filters that will select related data for track management systems

The following is an example of the code in Listing 3.4 for our HBase to get data from our SQL Server.

Listing 1. A Code Segment for HBase to get data from SQL Server.

```
… HADOOP_CLASSPATH=`/opt/hbase-0.92.1-
security/bin/hbase classpath`
${HADOOP_HOME}/bin/hadoop jar
/opt/hbase-0.92.1-security/hbase-0.92.1-
security.jar importtsv -
Dimporttsv.columns=HBASE_ROW_KEY,
asset:model, asset:serialnumber, test:device,
test:type, test:objective, test:operator, …,
event:time, algorithm:name, algorithm:date
'Device_test'
hdfs://RTanalytic…/hadoop/scrap/DeviceSQL4H
Base.Tag.txt …
```

4.2 Experimental Results:

In our experiments, the predictive accuracies obtained by the Rule Induction in our HD-MaLA Intelligent Agent have been compared to the predictive accuracies of four well-known, human-designed rule induction algorithms: namely ordered CN2, unordered CN2, Ripper, and C4.5Rules. The experimental results demonstrate that the HD-MaLA Agents are more competitive with these baseline algorithms and produce better rule induction algorithms (with higher predictive accuracy) than a grammar-based hill-climbing system.

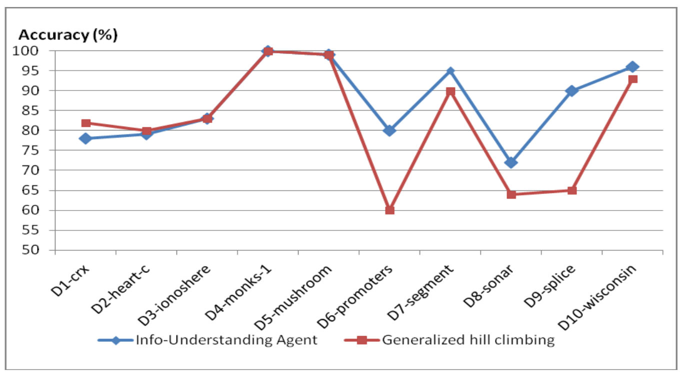


Figure 3. Predictive accuracies (%) between the HD-MaLA Intelligent Agent Rule Induction and GHC Rule Induction

We conducted experiments to evaluate the effectiveness of our framework in automatically designing rule induction algorithms, its sensitivity to parameters, and the effectiveness of our search method compared to other rule induction algorithms. As shown in Figure 3, one of our experiments illustrates that genetic programming based Information Understanding Agents predict accurately compared to the generalized hill climbing (GHC) algorithm using 10 benchmark datasets, D1 through D10.

5. FUTURE WORK

While the preliminary experiments are succesful, we further plan to conduct further performance evaluation experiments with respect to scalability and performance metrics related to decision speed, data to decision rate and optimized usage of resources compared to the related methods utlizing real time data.

6. CONCLUSION

Our HD-MaLA big data analytics framework will benefit government, military, industrial, and business applications, where high quality decision making capabilities are required. The experimental results demonstrate that our Big Data analytics framework based Intelligent Agents are more competitive with the baseline algorithms and produce better rule induction algorithms specifically higher predictive accuracy. In summary, HD-MaLA framework have the following advantages over classic and current data mining tools and products for big data analytics:

(a) Faster data-to-decision processing speed for real time military and business applications because our real time parallel and distributed computing algorithms are developed by using the Storm that do not require queries and data movement;
(b) Results in efficient computing because there is minimal data movement and communication traffic across networks (including virtual machines), and no raw data is moved for data mining around networks to significantly enhance computing efficiency and system security;
(c) Optimized usage of computing resources such as computing memories, bandwidth, and CPUs, since tiny codes of HD-MaLA Agents consume little computing resources; and
(d) Automated data seeking, processing, information understanding and presenting to significantly reduce decision maker's work load and dramatically increase data-to-decision rate, since HD-MaLA intelligent Agents automate most tasks or jobs that are completed by using Storm based scalable computing framework.

REFERENCES

[1] Apache Hadoop, http://hadoop.apache.org/ Accessed November, 2013
[2] Cloud MapReduce http://code.google.com/p/cloudmapreduce/ Accessed November, 2013
[3] Storm, http://storm-project.net/ Accessed November, 2013
[4] S4 http://incubator.apache.org/s4/ Accessed November, 2013
[5] Mahout http://mahout.apache.org/ Accessed November, 2013
[6] Brian Xu, Andrew Kostrzewski, "Networked Intelligent Agents for Distributed Decision Making." The Int'l Conf. on Artificial Intelligence (ICAI), P35-42, Las Vegas, 2007.
[7] X.F. Zha, "Artificial Intelligence and Integrated Intelligent Information Systems", , IGI Press, Hershey, PA USA, 2007
[8] Code Architecture (http://code.google.com /p/cloudmapreduce/wiki/CodeArchitecture Accessed October, 2013
[9] Brian Xu, Andrew Kostrzewski, "Integrated Intelligent Decision-Making and Information Discovery." The Int'l Conf. on Artificial Intelligence (ICAI), P 30-36, Las Vegas, NV, 2006.
[10] Brian Xu, A. Kostrzewski, P. LaMonica, R. Jurgens, "Net-Centric Rule Engines for Decision Support," The 2008 Int'l Conference on Artificial Intelligence, 2008, Las Vegas, Nevada.
[11] Brian Xu, "Integrated Link Analyzers and Associative Search Engines." The 2010 Int'l Conf. on Artificial Intelligence (ICAI), Las Vegas, NV, July 2010.
[12] Sean Owen, etc., "Mahout in Action", Manning Publications Co., 2012.
[13] I. Ari, E. Olmezogullari and F. Celebi, "Data stream analytics and mining in the cloud", 2012 IEEE 4th International Conference on Cloud Computing Technology and Science (CloudCom) pp. 857 -862
[14] S. Chaisiri et al.,"Workflow framework to support data analytics in cloud computing", 2012 IEEE 4th International Conference on Cloud Computing Technology and Science (CloudCom), December 2012, pp. 610-613
[15] L. Zhang, "Big Services Era: Global Trends of Cloud Computing and Big Data", IEEE Transactions on Services Computing, September 2012, pp. 467-468
[16] S. Guo et al., "A MapReduce-based System for Time-Based Big Data Analytics", 2012 IEEE International Conference on Cluster Computing (CLUSTER), pp. 72-80
[17] N. Leavitt, "Bringing Big Analytics to the Masses", Computer, Jan 2013, pp. 20-23

Weighted Semantic PageRank Using RDF Metadata on Hadoop

Hee-Gook Jun[1], Woo-Hyun Lee[1], Dong-Hyuk Im[2], Sang-Goo Lee[1], and Hyoung-Joo Kim[1]
[1] School of Computer Science and Engineering, Seoul National University, Seoul, Korea
[2] Dept. of Computer and Information Engineering, Hoseo University, Chungnam, Korea

Abstract - *PageRank, a representative link-based algorithm, evaluates the importance of Web pages based on the number of in-links each has. However, this feature may cause a problem in that pages with many in-links can be highly ranked regardless of their importance to the given query. Many methods have attempted to solve this problem by evaluating the weight of the links to stratify their importance. However, these methods have a limitation in that the weight of the links cannot be evaluated by their meaning directly owing to the hyperlink-based Web structure. We therefore propose a new approach to utilize the meaning of links directly by changing from a hyperlink-based Web structure to a semantic-link-based Web structure. In addition, we implemented the ranking method using the MapReduce framework to improve performance of semantic Big Data processing. The results of our experiment show that our approach outperforms the existing PageRank algorithm.*

Keywords: Big Data, MapReduce, Semantic Web, PageRank, RDF

1 Introduction

As the World Wide Web produces a greater amount of information over time, such information needs to be processed more effectively and efficiently to provide more accurate information to users. This issue has become a key challenge for Web-based information retrieval [8, 12, 19]. Since the 1990s, various methods dealing with the explosion of information on the Web have been studied in the field of Web information retrieval, including indexing, clustering, user interface methods, and ranking.

A page-ranking algorithm is an essential Web information retrieval method as the volume of results matched with a given query during a retrieval step is hard to be managed by users. Such an algorithm answers a given user query with a page list ranked by importance. The early page-ranking algorithm was a term-based ranking algorithm, whose criterion for evaluating the importance of pages is how many matched terms [3, 4] are contained on the page. After 1998, alternative ranking algorithms based on a linked relationship of pages [5, 11] were provided, and proved that link-based ranking algorithms perform better than term-based algorithms.

PageRank [5] is a representative link-based ranking algorithm. The authors of this algorithm assume that important pages are referred to by many other pages. Through this method, each page distributes its rank score to other pages they link to. Therefore, the more in-links a page has, the more important the page is. PageRank, however, which does not consider the semantics of links when computing their importance, may highly rank pages that only contain meaningless in-links.

Many algorithms have been suggested to tackle the above problem [18, 20, 26], and some have considered evaluating the weight of the links to adjust the propagation of their rank scores. In this way, if a page has many in-links with a small weight value, the page will have a lower value of importance. However, for a hyperlink-based Web structure, there are two significant limitations in evaluating the weight of a link. First, the hyperlink does not explain why pages are linked to other pages. Therefore, existing algorithms evaluate the weight of the links indirectly, such as by counting the number of links or analyzing other features out of links. Second, a page used as a unit of ranking is actually an object containing information rather than information itself. In other words, highly ranked pages that have high importance values owing to the presence of many in-links do not always contain important information, and may even contain meaningless information.

Moreover, the other problem of ranking is that ranking algorithms require a large space to store Web link structure and ranking values for every page. It is not easy for a single machine to compute large-scale data and to produce ranking results in a reasonable time. Hence searching an appropriate Big Data processing method for ranking is essential to deal with computing time and space problem.

In this paper, we propose the Weighted Semantic PageRank (WSPR) algorithm, which uses semantic links directly for a more accurate page ranking. We utilize RDF [15] metadata to create a semantic-link-based Web structure from a hyperlink-based Web structure, and utilize this semantic information as inputs for WSPR. Using semantic links in a semantic-link-based Web structure helps resolve the problem in determining the meaning of the links provided by a hyperlink-based Web structure. We can compute the rank scores in a semantic-link-based Web structure by evaluating the meaning of links directly. Furthermore, WSPR is able to reduce the possibility of ranking less important pages with high scores, as it uses RDF resources on the pages to compute their ranks, thus giving pages with less important resources a smaller ranking score. In addition, we implemented WSPR algorithm using the MapReduce framework [1], which is a more effective parallel distributed processing method for analyzing a large volume of Web data.

The contributions of this paper can be summarized as follows:

- We propose a ranking algorithm that computes the importance of pages more accurately. When the algorithm evaluates the weight of links, it considers the semantics of the link directly, and does not simply consider the number of links or use additional factors to estimate the meaning of links.

- The proposed algorithm prevents meaningless pages with many in-links to be mistaken for important pages. Resources, the semantic units of RDF instances, are used for computing the importance value instead of pages. Therefore, once a page receives a high importance value, the proposed algorithm guarantees that the page contains meaningful resources indicating important information.

- We developed WSPR system using MapReduce on Hadoop. This system has more computation capability for processing Big Data resources, enabling the proposed ranking algorithm to be utilized on the Web.

The remainder of this paper is organized as follows. In Section 2, we provide an overview of PageRank and Extended PageRank algorithms, focusing on the evaluation of the link weights. In Section 3, we introduce a semantic-link-based Web structure. In Section 4, we present our WSPR algorithm using the MapReduce framework in detail. Section 5 reports the results of our experiments used to evaluate the validity of our proposal. Finally, in Section 6, we offer some concluding remarks regarding the proposed research as well as some directions for future work.

2 Related Work

A page-ranking algorithm is an essential Web information retrieval method as the volume of results matched with a given query during a retrieval step is hard to be managed by users. Such an algorithm answers a given user query with a page list ranked by importance. The early page-ranking algorithm was a term-based ranking algorithm, whose criterion for evaluating the importance of pages is how many matched terms [3, 4] are contained on the page. After 1998, alternative ranking algorithms based on a linked relationship of pages [5, 11] were provided, and proved that link-based ranking algorithms perform better than term-based algorithms.

$$PR(r_i) = d \sum_{j \to i} \frac{1}{N_j} \cdot PR(r_j) + (1 - d) \qquad (1)$$

where d is a damping factor, which can be set between 0 and 1. This damping factor is used to resolve the rank sink problem caused by a cyclic linked or non-linked Web structure. The damping factor is usually set to 0.85. In PageRank, the PageRank value of a page is the sum of the PageRank values of pages that refer to this page. Each page equally distributes its PageRank value to pages it links to.

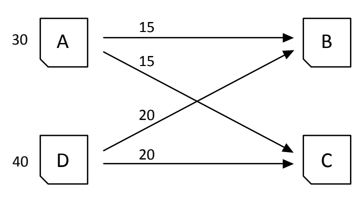
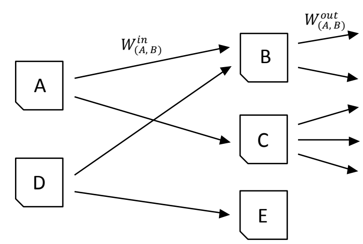


Fig. 1. A PageRank example.

In Figure 1, for example, page A, whose PageRank value is 30, assigns a PageRank of 15 to each of B and C. Similarly, page D assigns a PageRank of 20 to each of B and C. However, owing to its counting method, meaningless pages may be ranked highly by PageRank, which does not consider the meaning of links but only their number.

Weighted PageRank [26] is an alternative approach for avoiding a uniform distribution of rank values without proper consideration of the meaning of each linked relationship. Weighted PageRank evaluates the weight of the links to stratify the distribution of rank values (Figure 2). It computes the link weights using the proportions of in-links and out-links (Equation 2). However, this method is also based on the numbers of links, not their meaning. Furthermore, because a unit of ranking is page, it still exists for meaningless pages to be provided as important pages. Furthermore, because the method just estimates the importance of pages by their links, it does not always guarantee that a page actually contains important information.

$$W^{in}_{(v,u)} = \frac{I_u}{\sum_{p \in R(v)} I_p} \quad , \quad W^{out}_{(v,u)} = \frac{O_u}{\sum_{p \in R(v)} O_p} \qquad (2)$$



$$PR(B) = PR(A)\, W^{in}_{(A,B)}\, W^{out}_{(A,B)} + \cdots$$

Fig. 2. A Weighted PageRank example.

Weighted Page Content Rank [18] improves Weighted PageRank by adopting Web content mining, through which it not only computes the link weights, but also observes the correlation between a given query and the resulting pages. However, this method still computes the weight of the links based on their number, and requires an extra cost involved with the mining process. Other methods such as Topic-Sensitive PageRank [27] and personalized PageRank [28] compute page importance using query-biased and user-biased metric. But our purpose is to generate an integrated page ranking algorithm as well as analyze semantic Big Data. Thus we set the scope of our research focused on unbiased page importance evaluation algorithm.

3 Semantic-link-based Web Structure

Semantic markup languages have been developed for better processing of Web information. Three representative semantic markup languages are RDFa [16], Microformats [14], and Microdata [13]. In this paper, we mainly focus on RDFa when building a semantic-link-based Web structure. RDFa was published in 2004 and received W3C recommendation in 2008. RDFa is a method used to describe RDF notations in XHTML (Figure 3). Web documents with RDFa can be read by Web browsers and extracted to obtain semantic information through RDFa parsers. The extracted information is a form of RDF [15] metadata. RDF, which is a data model used to describe a set of knowledge, uses “triples” to express semantic relations among the knowledge set. A triple is composed of a subject, a predicate, and an object. This triple structure can be regarded as the unit of a graph dataset (Figure 4). Similar to RDF, RDFa is also a graph data model, and is more manageable for ranking algorithms than other semantic markup languages. Furthermore, this graph data model has an RDF predicate as a semantically labeled link, thus allowing the ranking algorithm to evaluate the weight of the links directly through their meaning.

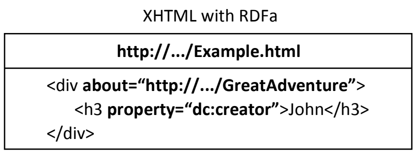


Fig. 3. An example RDFa annotation.

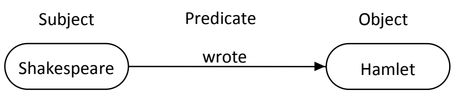


Fig. 4. An example RDF triple.

Several major sites have adopted RDFa which may help drive us toward the realization of the Semantic Web; Yahoo! and Google use RDFa for customizing their search results [21]; Facebook uses it for handling social data [22]; and Content Management Systems like Drupal and Wordpress, for semantic tagging [6]. Utilization methods of RDFa have also been provided; W3C has provided an RDFa distiller and a parser. In addition, RDFauthor [24] has provided an integrative approach for the management of RDFa data, and annotation systems [25, 7, 17, 10] have also been provided for various research fields.

Accordingly, it is clear that each of these methods is driving us closer to the existence of the Semantic Web. Thus, we consider the situation that pages contain semantic metadata using RDFa. If pages do not use RDFa notation for semantic metadata definition, we assume that these pages use other annotation method and use Information Extraction method to extract RDF format data.

4 Weighted Semantic PageRank

4.1 Proposed Architecture

Weighted Semantic PageRank (WSPR) system provides a new evaluation method that uses a semantic-link-based Web structure. It computes the weight of the links by evaluating their meaning directly. Four steps are used in this system (Figure 5). The first two steps change the environment from a hyperlink-based Web structure to a semantic-link-based Web structure. The other steps compute ranking values based on the structure constructed in the first two steps.

4.1.1 Semantic Information Extraction

As the first step of the WSPR algorithm, the system collects semantic information from the pages. Extracting semantic information is now much easier than before owing to our construction of a semantic-labeled Web structure described in the previous section. The system begins crawling through the Web using a parsing RDFa syntax. Figure 6 shows an example of RDF data extraction from a Web page containing an RDFa annotation. The WSPR algorithm uses a resource such as an object or a subject in an RDF triple as the unit of ranking. In other words, resources themselves are ranked, and the predicates are labeled links between resources.

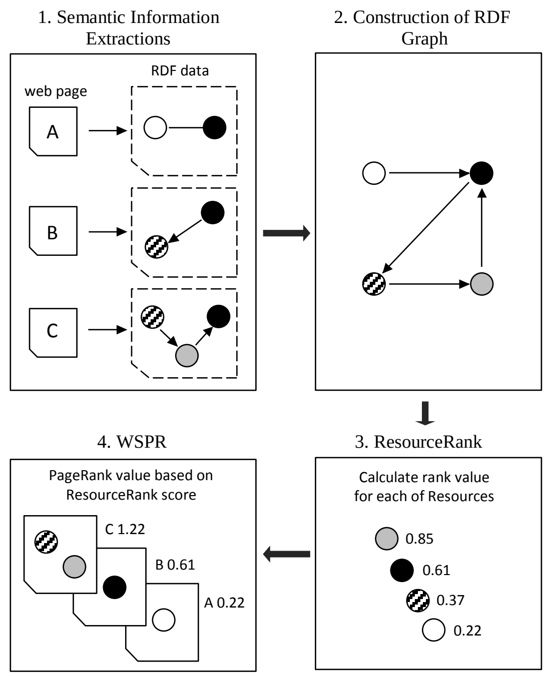


Fig. 5. Overview of the steps followed in the WSPR system.

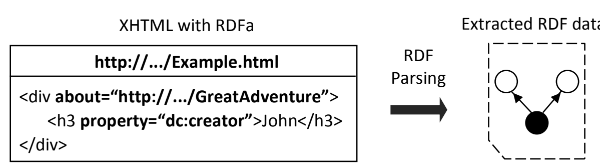


Fig. 6. RDF parsing of a Web page with an RDFa annotation.

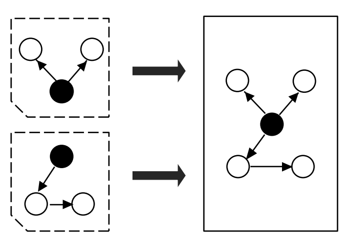


Fig. 7. Merging RDF triples with resources having the same URI.

4.1.2 Construction of RDF Graph

After the first step, the system then obtains a set of RDF data from the pages. A single series of RDF triples is insufficient for determining the rank value efficiently, and therefore multiple series must be interconnected together. Hence, this step uses a Uniform Resource Identifier (URI) as a key to find matched resources. For instance, supposing there are two triples to merge, as shown in Figure 7, the system checks the URI of their resources. When the system identifies that both triples have a resource with the same URI (the black nodes in Figure 7), the system merges the two triples into a single directed graph. In this way, this system creates a combined graph by merging all of the triples.

4.1.3 ResourceRank

In the third step, our system begins a ranking process called ResourceRank, which computes the ranking scores of the resources on the RDF graph built in the second step. ResourceRank can evaluate the weight of the links based on their own meaning since predicates labeled to contain semantic information are used to link the resources. This stratifies the distribution of rank values between linked resources based on the degree of their semantic relationships.

There are two types of methods, manual and automatic, used to evaluate the weight of the links [23]. We adopted the TF-IDF method to compute the weight of the links automatically. More specifically, the WSPR algorithm evaluates predicates instead of terms since it runs on an RDF graph. Therefore, it computes the Predicate Frequency (PF) as a Term Frequency. PF uses a function f which returns raw frequency of a predicate, and for normalization, the frequency divided by the maximum raw frequency of any predicate of the resource. IDF is also computed using predicates. The equations are as follows:

$$PF(p,r) = \frac{f(p,r)}{\max\{f(w,r): w \in r\}} \tag{3}$$

$$IDF(p,R) = log\frac{|R|}{|\{r \in R: p \in r\}|} \tag{4}$$

where p is a target predicate to compute the weight, r is a resource, and R is a set of resources.

Using PF-IDF, the value of a link weight is defined by using Equation 5.

$$weight(r_i,p) = PF(r_i,p) \times IDF(r_i,p) \tag{5}$$

Finally, ResourceRank equation takes on the form,

$$RR(r_i) = d \sum_{j \in outlink(i)} \frac{RR(r_j) \cdot weight(r_j,p)}{\sum_{j \in outlink(i)} weight(r_j\,,p)} + (1-d) \tag{6}$$

where RR(r_j) is the ResourRank value of a resource linked to resource ri, and is stratified based on its importance (weight) before being added to RR(r_i).

4.1.4 Weighted Semantic PageRank

The final step of this system is computing the PageRank value. In this step, the rank values of the pages are evaluated using the resource rank values calculated in the previous step. All resources originally contained on each page are in RDFa syntax forms. This means that the importance of a resource can be used to project the importance of the pages that contain this resource. That is, page importance is based on how many important resources, not how many in-links, a page has, unlike in previous ranking algorithms, which define page importance based on the latter criterion. This feature requires an important page with a greater PageRank value to contain important resources with meaningful information, and thus the probability that meaningless pages will be highly ranked is lower than in previous ranking algorithms.

Equation 7 shows the PageRank value using the ResourceRank values calculated in the previous step.

$$PageRank(p_i) = \sum_{r \in p_i} RR(r) \tag{7}$$

where RR(r) is the ResourceRank value of resource r, which is contained in page p_i. Thus, the PageRank value of page p_i is the summation of all ResourceRank values of the page pi resources.

4.2 MapReduce Algorithm

MapReduce methodology makes development of distributed and parallel processing more efficient. Google generated MapReduce framework, and Apache released Hadoop [2] - an open source implementation of Google's MapReduce framework. Hadoop has been extensively used on Big Data processing. On the MapReduce framework, researchers are able to concentrate on solving their own problems without having to manage distributed and parallel system directly.

A MapReduce job consists of map and reduce phases (Figure 8). In the map phase, input data is converted into key-value pairs. The key-value pairs are sent to the reduce phase by keys. In the reduce phase, data sets combined by key are processed for a specific purpose.

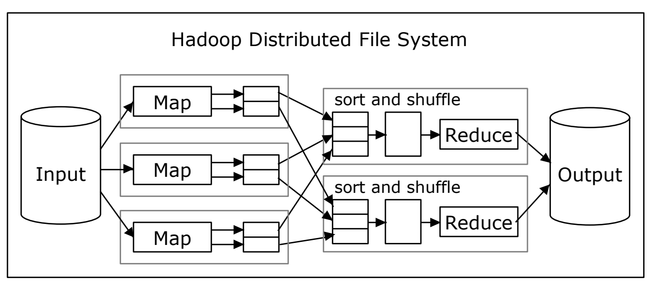


Fig 8. Overview of Hadoop MapReduce.

We implemented MapReduce version of WSPR in order to analyze large-scale semantic metadata. The WSPR MapReduce algorithm processes three jobs (Figure 9). The first job receives page information and their RDF metadata. The first job computes ResourceRank for each RDF resource until convergence, and the results of the first job are passed to the next job (Figure 10).

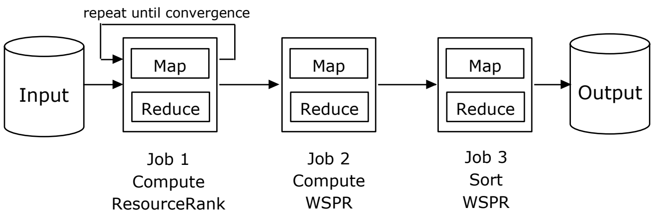


Fig. 9. WSPR MapReduce job framework.

```
class MAPPER
  method MAP(pageid i, page P)
    EMIT(pageid i, page P)   // Emit adjacency list
    for all pageid j ∈ P.AdjacencyList do
      r ← j.ResourceRank × j.LinkWeight
      EMIT(pageid j, r)        // Emit value for ResourceRank
    end

class REDUCER
  method REDUCE(pageid i, values [v1, v2, …])
    R ← ∅
    sum ← 0
    for all v ∈ values [v1, v2, …] do
      if IsResourceRankScore(v) then
        sum ← sum + v        // Sum of values for ResourceRank
      else
        R.AdjacencyList ← v   // Get adjacency list information
      end
    end
    R.ResourceRank ← sum × 0.85 + 0.15  // Compute rank
    EMIT(pageid i, page R)
```

Fig. 10. MapReduce Job 1: ResourceRank.

```
class MAPPER
  method MAP(pageid i, page P)
    EMIT(pageid i, P.resourceRank)

class REDUCER
  method REDUCE(pageid i, resourceRanks [r1, r2, …])
    R ← ∅
    sum ← 0
    for all r ∈ resourceRanks [r1, r2, …] do
      sum ← sum + r        // ResourceRank value summation
    end
    R.PageRank ← sum
    EMIT(pageid i, page R)
```

Fig. 11. MapReduce Job 2: WSPR.

```
class MAPPER
  method MAP(pageid i, page P)
    EMIT(P.PageRank, pageid i)  // Sort using Reduce function
```

Fig. 12. MapReduce Job 3: Ordering page by rank score.

In the second job, RDF resource information with ResourceRank score is used for computing WSPR score. RDF resource and RDF ResourceRank score pairs are grouped into pages each resource belongs to. WSPR score of each page is computed by summing up the group of ResourceRank scores assigned to each page (Figure 11).

The third job takes intermediate ranking information from the previous job as input data. Finally, the third job sorts pages by WSPR score and outputs the ranking result (Figure 12).

5 Experimental Evaluation

5.1 The Setup

The physical Hadoop cluster for the experiments comprises one master node and eleven slave nodes. Each node has 3.1 GHz quad-core CPU, 4GB memory, and 2TB hard disk. The operating system is 32-bit Ubuntu 12.04.2, the java version is 1.6.0_26, and the Hadoop version is 1.2.1.

As a source of Web data, we used 80,000 WikiPedia [9] web pages and extracted 500,000 RDF metadata from infobox tables in the WikiPedia pages.

5.2 Results

We evaluate WSPR and other systems according to precision, recall and f-measure. In Figure 13, the solid line indicates the results of WSPR, the dashed line with filled triangle is those of Weighted PageRank (WPR), and dashed line with cross is PageRank (PR) result. The result shows that WSPR has higher evaluation values than the others. This means WSPR provides less false positive and false negative ranking results. Similar conclusion can be drawn from Table 1, which shows the comparison among NDCG [29] of PR, WPR, and WSPR. We see that the results of WSPR attain higher values than those of the others.

Table 1. NDCG@k results for the test query

NDCG@k	PR	WPR	WSPR
NDCG@5	0.8765	0.9838	0.9931
NDCG@8	0.8824	0.9469	0.9748
NDCG@10	0.8866	0.9389	0.9732

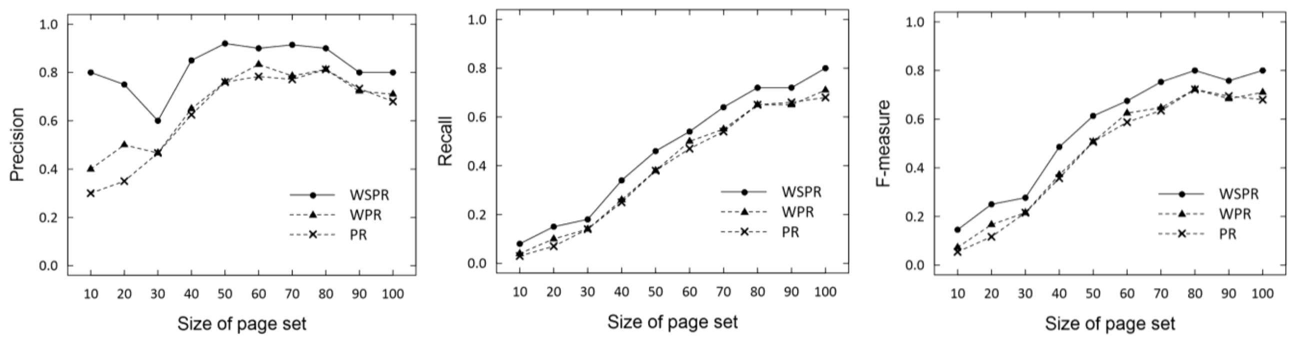


Fig. 13. Precision, Recall, and F-measure of PR, WPR, and WSPR for varying number of pages.

Table 2 shows a more detailed view of query results on literatures. In the ResourceRank stage, the third step of WSPR, the page on Macmillan has two resources: “Macmillan” and “Publishing company”. The ResourceRank scores of these two resources are 1.118 and 0.429, respectively. On the other hand, the page on the United States has one resource, “United State,” the ResourceRank score of which is 1.272. Although “United State” has the highest ResourceRank value among the three resources, the page on Macmillan has a higher WSPR score than the page of the United States (Table 3).

It is natural for human to choose the page on Macmillan as the most related page to the given query, since Macmillan is a publishing company, while the page on the United States appears irrelevant to be chosen as a related page. This shows that the result of WSPR takes semantic meanings of the pages into account.

Table 2. ResourceRank related within pages

RDF Resource	ResourceRank Score
“United State”	1.272
“Macmillan”	1.118
“Publishing company”	0.429

Table 3. Summary of ResourceRank used to compute WSPR

Page	RDF Resource (ResourceRank Score)	WSPR Score
Macmillan	“Publishing company” (0.429) “Macmillan” (1.118)	1.547
United States	“United State” (1.272)	1.272

Next we measured the processing time of semantic Big Data analysis. Figure 14 shows the execution time for the experiments. Each result with different data size supports linear growth in processing time rather than exponential growth. Thus, the results indicate that WSPR implemented using the MapReduce framework has two benefits in processing Big Data. First, it enables computation of large-scale semantic data. Second, the computation of the data takes relatively small amount of time compare to the other algorithms.

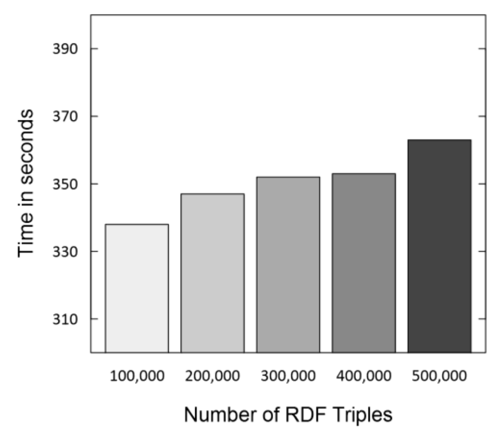


Fig. 14. MapReduce execution time

6 Conclusions

An RDF model can define concepts through the use of triples, which have a semantic link structure. In this paper, we utilized this feature to resolve a problem with PageRank in which the meaning of links used to compute importance cannot be properly evaluated.

Using a new ranking method that can be used to evaluate importance based on how many important resources Web pages have, WSPR provides more semantically relevant ranking results than other ranking methods. Therefore, once a page is ranked highly by WSPR, the page is guaranteed to contain important information related to the given query as WSPR ranks pages based on how many important resources, not how many in-links, the pages have, avoiding meaningless pages to be scored highly, which is a problem with other ranking methods.

Furthermore, we have adopted MapReduce framework to compute WSPR. Performance evaluations show that parallel and distributed processing on Hadoop is an effective way for semantic Big Data analysis.

Further research will be conducted using an automatic RDFa annotator. This will enable the WSPR algorithm to use both Web pages without semantic metadata and semantically annotated Web pages for computing a semantic rank score. We expect this to improve the adoptability of WSPR across the World Wide Web.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIP) (No. 20110030812), by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIP) (No. 20110017480), and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A1002236).

References

[1] J. Dean and S. Ghemawat. “Mapreduce: Simplified data processing on large clusters,” In Proceedings of the 6th USENIX Symposium on Opearting Systems Design & Implementation (OSDI), pp. 137-150, 2004.

[2] Hadoop, Available from: http://hadoop.apache.org [Accessed: 12 March 2014].

[3] R. J. Bayardo, Y. Ma, and R. Srikant, “Scaling up all Pairs Similarity Search.” In Proceedings of the 16th international conference on World Wide Web, pp. 131-140, 2007.

[4] G. Salton and C. Buckley, “Term-weighting approaches in automatic text retrieval,” Information Processing and Management, Vol. 24, No. 5, pp. 513-523, 1988.

[5] S. Brin and L. Page, “The Anatomy of a Large-scale Hypertextual Web Search Engine,” Computer Networks and ISDN Systems, Vol. 30, No. 1-7, pp. 107-117, 1998.

[6] S. Corlosquet, R. Cyganiak, A. Polleres, and S. Decker, “RDFa in Drupal: Bringing cheese to the web of data,” In Proceedings of the 5th Workshop on Scripting and Development for the Semantic Web at ESWC, 2009.

[7] M. Duma, “RDFa Editor for Ontological Annotation,” In Proceedings of the Student Research Workshop associated with RANLP, pp. 54-59, 2011.

[8] V. N. Gudivada, V. V. Raghavan, W. I. Grosky, and R. Kasanagottu, R, “Information Retrieval on the World Wide Web,” IEEE Internet Computing, Vol. 1, No. 5, pp. 58-68, 1997.

[9] Wikipedia, Available from: http://en.wikipedia.org/ [Accessed: 12 March 2014].

[10] A. Khalili, S. Auer, and D. Hladky, “The RDFa Content Editor - From WYSIWYG to WYSIWYM,” Proceedings of the IEEE Signature Conference on Computers, Software, and Applications, COMPSAC, 2012.

[11] J. M. Kleinberg, “Authoritative Sources in a Hyperlinked Environment,” Journal of the ACM, Vol. 46, No. 5, pp. 604-632, 1999.

[12] M. Kobayasha and K. Takeda, “ Information Retrieval on the Web,” ACM Computing Surveys, Vol. 32, No. 2, pp. 144-173, 2000.

[13] W3C Working Group, "HTML Microdata," Available from: http://www.w3.org/TR/2011/WD-microdata-20110405/ [Accessed: 12 March 2014].

[14] R. Khare, “Microformats: The Next (Small) Thing on the Semantic Web?,” Journal IEEE Internet Computing archive, Vol. 10, No. 1, pp. 68-75, 2006.

[15] RDF Working Group, “Resource Description Framework,” Available from: http://www.w3.org/RDF/ [Accessed: 12 March 2014].

[16] W3C Working Group, "RDFa Core 1.1 - Second Edition," Available from: http://www.w3.org/TR/rdfa-syntax/ [Accessed: 12 March 2014].

[17] M. Samwald, E. Lim, P. Masiar, L. Marenco, H. Chen, T. Morse, P. Mutalik, G. Shepherd, P. Miller, and K. Cheung, ”Entrez Neuron RDFa: A Pragmatic Semantic Web Application for Data Integration in Neuroscience Research,” In Proceedings of the International Conference of the European Federation for Medical Informatics, pp. 317-321 , 2009.

[18] P. Sharma, D. Tyagi, and P. Bhadana, “Weighted Page Content Rank for Ordering Web Search Result,” International Journal of Engineering Science and Technology, Vol. 2, No. 12, pp. 7301-7310, 2010.

[19] A. Singhal, “Modern information retrieval: A brief overview,” Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, Vol. 24, No. 4, pp. 35-43, 2001.

[20] K. Stein and C. Hess, “Information retrieval in trust-enhanced document networks,” In Proceedings of European Web Mining Forum (EMWF) 2005, and Knowledge Discovery and Ontologies (KDO) 2005.

[21] T. Steiner, R. Troncy, and M. Hausenblas, “ How Google is using Linked Data Today and Vision For Tomorrow,” Future Internet Assembly, Ghent, Belgium, 2010.

[22] The Open Graph Protocol, Available from: http://ogp.me [Accessed: 12 March 2014].

[23] N. Toupikov, J. Umbrich, R. Delbru, M. Hausenblas, and G. Tummarello, “DING! Dataset Ranking Using Formal Descriptions” In Proceedings of WWW 2009 Workshop on Linked Data on the Web (LDOW 2009), Madrid, Spain, 2009.

[24] S. Tramp, N. Heino, S. Auer, and P. Frischmuth, “RDFauthor: Employing RDFa for collaborative knowledge engineering,” In Proceedings of Cimiano, P., Pinto, H.S. (eds.) EKAW 2010. LNCS, Vol. 6317, pp. 90-104. Springer, Heidelberg , 2010.

[25] R. D. Virgilio, F. Frasincar, W. Hop, and S. Lachner, “A Reverse Engineering Approach for Automatic Annotation of Web Pages,” Multimedia Tools and Applications, Published online, 2011.

[26] W. Xing and A. Ghorbani, “Weighted PageRank Algorithm,” In Proceedings of Second Annual Conference on Communication Networks and Services Research CNSR 2004, Fredericton, N.B., Canada, pp. 305-314, 2004.

[27] T. H. Haveliwala, "Topic-sensitive PageRank," In Proceedings of the 11th international conference on World Wide Web, pp. 517-526, 2002.

[28] G. Jeh and J. Widom, "Scaling personalized web search," In Proceedings of the 12th international conference on World Wide Web, pp. 271-279, 2003.

[29] K. Järvelin and J. Kekäläinen, “Cumulated gain-based evaluation of IR techniques,” ACM Transactions on Information Systems, Vol. 20, No. 4, pp. 422-446, 2002.

Anonymization Infrastructure for Secondary Use of Data

Yuichi Nakamura, Kanae Matsui, Hiroaki Nishi
Graduate School of Science and Technology, Keio University, Japan
{nakamura, matsui}@west.sd.keio.ac.jp, west@sd.keio.ac.jp

Abstract— **Data containing sensitive and personal information is critical to the functioning of institutions in numerous fields, such as medical, transportation, and government. Moreover, these types of data are gaining value for secondary uses, such as market research, estimation of a route of infection, and traffic pattern analyses. From a privacy preservation viewpoint, publishing the raw data may raise significant issues because of the sensitive nature of the relevant data. Therefore, an infrastructure for publishing sensitive data while protecting privacy is required, to enable secondary use of the data. In this paper, we propose an infrastructure that supports secondary use of sensitive data in a secure manner. The proposed infrastructure preserves privacy by utilizing anonymization to publish the data; furthermore, the anonymizing process employs both a publishing rule and a request rule, thereby enhancing security. Additionally, a format for publishing datasets and their privacy-preserving rules is proposed, and is termed the XML-based Anonymize Sheets (XAS). The publishing organization and the secondary consumer of the data can designate publishing permissions and requests by utilizing XAS. The proposed infrastructure prevents additional leaks of sensitive information by utilizing the previously anonymized data as a publishing history.**

Keywords— secondary use of data, anonymization infrastructure, XAS

I. Introduction

Various institutions such as medical facilities, transportation facilities, and government agencies must manage large amounts of data, which may include customer information, medical records, and transaction information. This data, commonly stored in electronic form, often contains sensitive personal information. These types of data are useful, and frequently necessary, to facilitate the provision of advanced services. However, stored data may contain a considerable amount of information about individuals. This may include basic information such as age, address as well as more sensitive items such as financial data, medical records, personal preferences and history of behavior. The data contain sensitive information that organizations must protect from unauthorized use.

Recently, a movement known as Linked Open Data (LOD) has attempted to facilitate sharing of these type of data, with a goal of increasing its value. As an outgrowth of LOD, another movement known as Open Government encourages citizens to monitor government activities by publishing certain government information. For example, the United States government publishes information including economic conditions and citizens' activities on Data.gov [1].

Secondary uses of data, including location information recorded by mobile phones and data from electricity smart meters, are under consideration in Japan. The location data of mobile phones will reveal the daily travels of their users. For example, some car navigation systems utilize mobile phones to connect to datacenters, and therefore, it can obtain the car's location and other relevant data. The primary purposes of these data are to track the requirements of car's maintenance and to facilitate road services for drivers. By analyzing the data, it is possible to obtain the driving speed and location of the car. In addition, analysis of this data can identify intersections where drivers frequently brake in a sudden manner. Utilizing this information, a road maintenance squad can check the intersection, where they may identify problems such as hidden or missing signs. Data from a smart meter can provide information about the daily activities of a household. Remote observation services that monitor elderly parents attract significant attention in an aging society. These examples demonstrate that the secondary use of data can potentially create new services while enhancing the data's value. From numerous viewpoints, the secondary use of data is under consideration, and its demand is increasing.

In equal measure, this secondary use of data can result in privacy problems. In the previous examples, the location data produced by a smartphone reveals the user's location at a given time. The amount of electricity usage recorded by smart meters may reveal excessive power consumption by a household, potentially revealing their high-income status. Moreover, it is simple to publish sensitive data utilizing the Internet without proper regard to privacy. If access to this information is not adequately restricted, it may promptly result in its unauthorized use. Aside from its usefulness, publishing the data may result in the infringement of privacy rights. Therefore, techniques for publishing the data while simultaneously protecting privacy are required for the safe secondary use of the data.

To address this problem, Privacy-Preserving Data Mining (PPDM) [2, 3] and Privacy-Preserving Data Publishing (PPDP) [4, 5] are proposed. These techniques have the ability to mine or publish the data without personally identifiable information, thereby protecting privacy. Anonymization is a practical technology that supports privacy protection[5]. Anonymization technology can adjust to different privacy protection levels, thus providing flexible privacy protection. A considerable variety of studies on this technique have been performed owing

TABLE 1 MEDICAL RECORD ($k = 2$)

	Birth	Gender	ID	Problem
$t1$	1970	male	121	cold
$t2$	1970	male	121	obesity
$t3$	1970	male	121	diabetes
$t4$	1980	female	121	diabetes
$t5$	1980	female	121	obesity
$t6$	1981	male	125	diabetes
$t7$	1981	male	125	cold

TABLE 2 ANONYMIZED MEDICAL RECORD ($k = 3$)

	Birth	Gender	ID	Problem
$t1$	1970	male	121	cold
$t2$	1970	male	121	obesity
$t3$	1970	male	121	diabetes
$t4$	198*	human	12*	diabetes
$t5$	198*	human	12*	obesity
$t6$	198*	human	12*	diabetes
$t7$	198*	human	12*	cold

TABLE 3 ANONYMIZED MEDICAL RECORD ($k = 3$)

	Birth	Gender	ID	Problem	
$t1$	1970	female	121	cold	Alice
$t2$	1970	female	121	cold	
$t3$	1970	female	121	cold	
$t4$	198*	human	12*	poor circulation	
$t5$	198*	human	12*	poor circulation	
t6	198*	human	12*	headache	Bob
t7	198*	human	12*	headache	

to its high versatility. It is one of the most preeminent privacy protection technologies in current use. Generalization and deletion of the data are necessary to prevent privacy infringements. However, they reduce the value of the data. As a result, there is a trade-off relationship between privacy protection and the utilization of the data.

Although techniques such as PPDM and PPDP have been investigated in numerous studies, a method of securely publishing the data to enable secondary use has not been definitively established. Furthermore, after calculating and publishing anonymized data from a data source, another anonymized data set, calculated and published from the same source may cause a privacy information leak if an unauthorized person can access both sets of anonymized data. When calculating and publishing anonymized data, it is necessary to consider all of the previously published data from the same source.

Considering these issues, it is crucial to establish a clear suggestion of technological guidance, an infrastructure, and a technical standard of protocols for the secondary use of data. The development of the protocol and infrastructure is especially important to its development. It will facilitate collaboration between organizations that produce the data and the companies that require the data for secondary use, and thus increase their data publishing activity. It will develop the market for secondary uses of data in conjunction with advanced services such as market research, estimation of a route of infection, and traffic pattern analysis. Moreover, it will reduce the utilization costs for both providers and consumers of secondary use data, owing to the unification of data processing procedures.

In this study, we propose a data-publishing infrastructure for secondary data use in conjunction with privacy protection by utilizing anonymization. In addition, we propose a protocol and XML-based data format for the proposed infrastructure. The infrastructure prevents further leaks of private information by employing the previously anonymized data as a publishing history.

This paper is arranged as follows. Features of anonymization and the associated privacy protection levels are described in Section 2. The design of the data format, protocol, and infrastructure for secondary data use is proposed in Section 3. The implemented mechanism is explained in Section 4. The evaluations of the proposed infrastructure are described in Section 5. Finally, we conclude the paper in Section 6.

II. ANONYMIZATION

Anonymization is one of the methods included in PPDM and PPDP. This method protects sensitive information by masking or generalizing the sensitive data. In addition,, it allows the adjustment of the privacy protection level. There are several generalization methods available for anonymization. In the following paragraphs, two relatively basic and frequently referenced generalization methods, k -anonymity and l -diversity are explained.

A. k-anonymity

K -anonymity is one of the methods utilized for generalization,[6] and it is the base of l-diversity. Further explanation of this method will incorporate the various definitions listed below.

(i) Data table

In this paper, a data list similar to a database table is termed a “data table.” Its column is termed an “attribute.” Address, birth, and gender are examples of attributes. One group of data corresponding to person or group of people is termed a “data set” and one data set is termed a “tuple.”

(ii) Attribute

An attribute among a group of related attributes that can identify a corresponding person by itself, such as name or unique ID, is termed an “identifier,” and others that cannot identify a group on their own, however, it can provide identification when combined with other attributes, such as illness, birth, gender, is termed a “quasi-identifier.”

(iii) Sensitive attribute:

A significant attribute for secondary use is termed a “sensitive attribute,” which can be selected from attributes that are not identifiers. The method will exclude this attribute from masking or generalization by anonymization. Furthermore, tuple groups that have the same quasi-identifier values are termed “q*-block.”

The definition of k-anonymity is as follows: “*In each q**-block in the data table, at least k tuples are included.*”

TABLE 1 represents an example of a medical records data table. In this table, the sensitive attribute is “Problem” and the quasi-identifiers are “Birth,” “Gender,” and “ID.” The data consists of a $t1$~$t3$ q*-block, a $t4$, $t5$ q*-block, and a $t6$, $t7$ q*-block. It represents $k = 2$. Even if an attacker attempts to ascertain a specific individual’s problem and has already obtained the individual’s quasi-identifier, the attacker can narrow the results down to only two tuples. **TABLE 2** indicates that the anonymization results from **TABLE 1** are $k = 3$. The results displayed in this table demonstrate that anonymization methods provide the required privacy protection level, utilizing masking or generalization.

As displayed in these tables, the masking or generalization processes prevent an attacker from identifying a specific person. There are several algorithms for calculating masking or generalization. The most popular algorithm is the heuristic searching method, utilizing double-nested loops.

B. l-diversity

l-diversity is a method designed to protect the privacy of data [7]. This method considers the diversity of sensitive attributes, and it is therefore different from k-anonymity.

The definition of l-diversity is as follows: “*In all q**-blocks in a data table, there are at least l different sensitive attributes.*”

Researchers designed this method to provide protection from the following attacks.

(i) Homogeneity attack:

TABLE 3 is additional example of a medical record data table. In this case, if an attacker has acquired Alice’s quasi-identifier, the attacker can read Alice’s problem from this table, because no diversity exists for the sensitive attributes in the q*-block.

(ii) Background knowledge attack:

Although the $t4$~$t7$ q*-block in **TABLE 3** has a diversity of sensitive attributes, if the probability of poor circulation is very low for males and an attacker is aware of that, the attacker can read Bob’s problem from **TABLE 3**.

l-diversity provides more security than k-anonymity for preserving privacy. However, the calculation cost of l-diversity is higher than k-anonymity.

III. SECONDARY USE INFRASTRUCTURE

The demand for the secondary use of the data such as medical records is increasing, because it may enable the estimation of infection routes. However, medical data frequently includes sensitive and private information. The medical data providers should define the anonymization methods and the related privacy protection levels when publishing the data. In addition, when the data provider permits several methods of anonymization, the consumers of the data must select a method that matches their requirements. Moreover, consumers of the anonymized data should avoid obtaining private data that exceeds their requirements, including situations where the data provider permits the lower protection level and thus provides the private data. Therefore, the anonymization data infrastructure should provide a method to define anonymization methods and protection levels that fulfill the requirements for both data providers and data consumers.

To meet these requirements, data publishing with anonymization is required. However, PPDP utilizing anonymization has numerous problems. One of the problems is

TABLE 4 MEDICAL RECORD ($\boldsymbol{k = 1}$)

	Birth	Gender	Problem
$t1$	1970	male	cold
$t2$	1970	male	obesity
$t3$	1970	male	diabetes
$t4$	1981	male	diabetes
$t5$	1981	female	obesity
$t6$	1982	female	diabetes
$t7$	1982	female	cold

TABLE 5 ANONYMIZED MEDICAL RECORD ($\boldsymbol{k = 2}$)

	Birth	Gender	Problem
$t1$	1970	male	cold
$t2$	1970	male	obesity
$t3$	1970	male	diabetes
$t4$	1981	human	diabetes
$t5$	1981	human	obesity
$t6$	1982	female	diabetes
$t7$	1982	female	cold

TABLE 6 ANONYMIZED MEDICAL RECORD (1) ($\boldsymbol{k = 3}$)

	Birth	Gender	Problem
$t1$	19*	male	cold
$t2$	19*	male	obesity
$t3$	19*	male	diabetes
$t4$	19*	male	diabetes
$t5$	198*	female	obesity
$t6$	198*	female	diabetes
$t7$	198*	female	cold

TABLE 7 ANONYMIZED MEDICAL RECORD (2) ($\boldsymbol{k = 3}$)

	Birth	Gender	Problem
$t1$	1970	male	cold
$t2$	1970	male	obesity
$t3$	1970	male	diabetes
$t4$	198*	human	diabetes
$t5$	198*	human	obesity
$t6$	198*	human	diabetes
$t7$	198*	human	cold

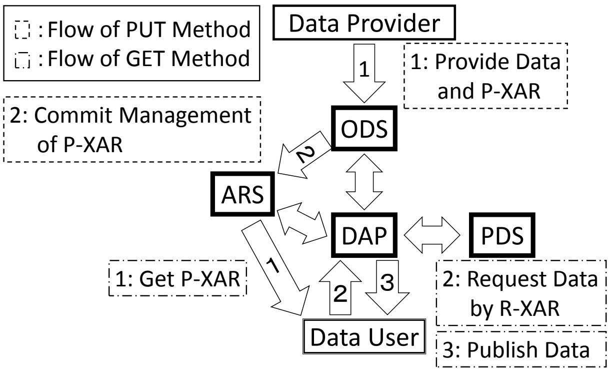


Fig. 1 Overview of proposed infrastructure

that no protocols and formats currently exist to enable secure data publishing, as described in the introduction. The other is loss of anonymity by publishing the same data multiple times. **TABLE 4** is an example of a medical record data table. **TABLE 5** is an anonymized $k = 2$ data table with data from **TABLE 4**, and **TABLE 6** is another anonymized $k = 3$ data table with data from **TABLE 4**. In this case, those who can obtain both the anonymized data of $k = 2$ and $k = 3$ can obtain the $k = 1$ data, including situations where the data provider did not permit the publishing of $k = 1$ data. This results in leak of privacy information. One cause of this problem is that previously published data is not referenced in the anonymization process; as a result the coherence between the $k = 2$ and $k = 3$ data was severed. **TABLE 7** is another example of a $k = 3$ data table. Utilizing **TABLE 7** instead of **TABLE 6** avoids the problem described above. **TABLE 7** was generated by anonymizing **TABLE 5** instead of anonymizing **TABLE 4**, to maintain coherency in masking and generalization. The proposed anonymizing process can prevent further leaks of privacy information.

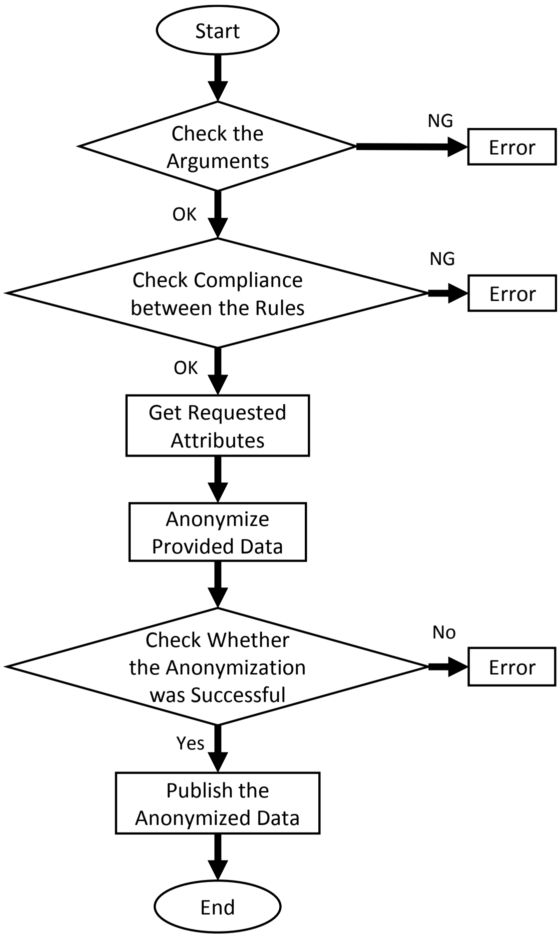


Fig. 2 Flowchart of the Implemented Mechanism

To address these problems, we proposed a data-publishing infrastructure. It manages the previously published data for the anonymization without the loss of anonymity, and provides safe secondary use and anonymization. For encryption technology, it utilizes Public Key Infrastructure (PKI). Certificate Authority serves a function as an authorized organization for certifying the public key of servers on the Internet. For this discussion, the anonymization technology and proposed infrastructure can be associated with the encryption technology and PKI, respectively.

A. Design of secondary use infrastructure

The proposed infrastructure can be divided into four organizations as follows.

(i) Original data storeroom organization (ODS)

This organization manages data provided by the data folder. The data folder is considered the data provider when the data is managed by ODS. When providing data to ODS, the data folder prepares data for publishing and provides an allowance rule by utilizing a specially designed format. This format is termed XML-based Anonymization Sheets (XAS). The details of XAS are described in the following section. Publishing rule descriptions utilize a subset of XAS, termed XML-based Anonymization Rules (XAR). The data folder generates data as D-XAS and the publishing rules (P-XAR) correspond to the D-XAS. D-XAS should include the link to the P-XAR. ODS should be responsible for maintaining the original data written as D-XAS in a secure manner. This data registration process is based on the PUT method.

(ii) Anonymizing rules storeroom organization (ARS)

This organization manages P-XAR. P-XAR will be openly published for users who need to access anonymized data based on the original data. P-XARs stored in the ARS can exhibit data when it is available for its secondary use. A P-XAR is stored by utilizing a PUT method issued by ODS.

(iii) Data anonymizing and publishing organization (DAP)

This organization anonymizes the original data (D-XAS) based on a publishing rule (P-XAR) and a request rule (R-XAR). A secondary use data consumer generates an R-XAR and provides it to the DAP. An R-XAR contains relevant information for D-XASs such as a URL, the requested anonymization method, its privacy level and anonymization range required to obtain the data for secondary use. The DAP receives the header of the requested D-XAS to access the link of the R-XAR. This header information does not include data. This header information is also described by using an XAR termed H-XAR; the DAP verifies its compliance by checking with the R-XAR and P-XAR requested from the ARS, according to the H-XAR. In this process, a user utilizes a GET method in conjunction with the R-XAR option. If it returns a compliance error, the user receives an appropriate error message. This message utilizes the HTTP error message protocol. If no error occurs, DAP issues a GET message to obtain the D-XAS from the ODS, and issues a subsequent GET message to receive the published XAS (P-XAS) from the PDS. The PDS is described in the following paragraph (iv). The DAP generates P-XASs as anonymized data, and the response from the R-XAR of the user. The user receives the anonymized data resulting from the GET method. Finally, the DAP stores the generated P-XAS issues by utilizing the PUSH method. This P-XAS is utilized to prevent further privacy leaks.

(iv) Published data storeroom organization (PDS)

This organization manages data previously published by the DAP as P-XASs. It may store all anonymized data generated by the DAP. However, to optimize data storage capacity, it is sufficient for the PDS to store only one P-XAS as anonymized data for each D-XAS, according to the one-direction anonymization policy. When generating P-XASs from D-XASs according to the requested R-XAR, it is sufficient to generate P-XASs according to the R-XAR, and store the P-XAS to the PDS. However, when generating another P-XAS from the same D-XAS according to another R-XAR, the DAP should obtain all P-XASs related to the D-XAS from the PDS. The DAP should consider all of these P-XASs when generating new P-XASs to observe P-XARs. Therefore, we propose one-directional anonymization to avoid this process. The process is as follows.

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-anonymize type="text/xas" href="p-xar.xas"?>
<list>
 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:v="http://www.w3.org/2006/vcard/ns#">
  <v:Kind rdf:about = "http://foo.com/me/hogehoge" >
   <v:fn>Hoge Foo</v:fn>
   <v:bday>1980-01-01</v:bday>
   <v:hasTelephone>
    <rdf:Description>
     <rdf:value>+81-45-566-1454</rdf:value>
     <rdf:type rdf:resource="http://www.w3.org/2006/vcard/ns#Work"/>
     <rdf:type rdf:resource="http://www.w3.org/2006/vcard/ns#Voice"/>
    </rdf:Description>
   </v:hasTelephone>
   <v:hasAddress>
    <rdf:Description>
     <v:street-address>123-45 Hoge Village</v:street-address>
     <v:locality>FooCity</v:locality>
     <v:postal-code>5555</v:postal-code>
     <v:country-name>Japan</v:country-name>
    </rdf:Description>
   </v:hasAddress>
  </v:Kind>
 </rdf:RDF>
 <OfficeScale>100ha</OfficeScale>
 <PowerConsumption>10kWh</PowerConsumption>

 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:v="http://www.w3.org/2006/vcard/ns#">
  <v:Kind rdf:about = "http://foo.com/me/db" >
```

Fig. 3 D-XAS Example (Extract)

(i) The DAP generates P-XASs according to P-XARs, instead of R-XARs, and stores it in the PDS. Therefore, the PDS stores the anonymized data, and it is anonymized according to the declared level in P-XAR. This P-XAS is not sent to the users if the requested level in the R-XAR is higher than the level in the P-XAR; this indicates the $k$value is larger than that of the P-XAR in k-anonymity.

(ii) DAP generates P-XASs according to the R-XARs. In this generation, the DAP only uses the first P-XAS generated from the P-XAR. DAP generalizes new P-XASs by adding "wild cards" as masking from the initial P-XAS. The DAP does not remove any of the "wild cards" provided as masking in the first P-XAS. Therefore, a one-directional anonymizing process should be considered.

(iii) The DAP can generate any type of P-XAS that satisfies both the R-XAR and the P-XAR by following the process described in (i) and (ii). In a scenario where k-anonymity and l-diversity are mixed, it is sufficient to generate a P-XAS that has a lower anonymization level than k-anonymity and l-diversity. For example, assume that 3-anonymity and 3-diversity are permitted in P-XARs, and 4-diversity is requested by R-XAR. In this case, DAP generates the initial P-XAR by utilizing 3-anonymity. The DAP can generate any type of P-XAR by utilizing the initial P-XAR, according to the one-directional anonymizing process.

To enable the data transfer between these organizations, data providers and data consumers will utilize SSL and PKI if they transfer the data over the Internet. In the following discussions, four organizations are exhibited in order to clarify each role. It is possible to merge some of them into a single organization.

Fig. 1 represents proposed organizational structure and data connections between the organizations.

IV. XML-BASED ANONYMIZE SHEETS (XAS)

We propose XML-based Anonymization Sheets (XAS) as a format to define the rules and data descriptions. To distinguish the rules from the data, XML-based Anonymization Rules (XAR) are also proposed as a subset of XAS. XAS and XAR differ because XAR does not contain data as contents. All transactions in the proposed infrastructure utilize the XAS and its subset, XAR. XAS is designed according to Extensible Markup Language (XML). **Fig. 3** lists an example of D-XAS. It includes the information to enable anonymization, including combinations of the sensitive attribute names and quasi-identifiers, permitted anonymization methods and levels, and data attributes such as created date, updated date and history, ownership, copyrights, comments, and others. **Fig. 4** lists an example of a P-XAR. It does not contain raw data; it only declares the required anonymization methods and levels. To enable masking or generalization processes, it can define the delimiter for distinguishing data sections. In this example, "BirthDay" is split utilizing the '-' character. During the anonymizing process, the character is used to define the generalization boundary. If the data employs a general and standardized format, for example, BirthDay should be separated by '-,' it can generalize the data entry by referring to the default rule. As an additional feature, the data provider may

```
<?xml version="1.0" encoding="utf-8"?>
<anonymize>
 <head>
  <publishacceptance sensitive="divisional" quasi="divisional" />
  <firstdatasetposition>
   <list>
    <rdf:RDF />
   </list>
  </firstdatasetposition>
  <sensitive type="k(>=3), l(>=2)">
   <rdf:RDF>
    <v:Kind>
     <v:hasTelephone>
      <rdf:Description>
       <rdf:type number="2" />
      </rdf:Description>
     </v:hasTelephone>
    </v:Kind>
   </rdf:RDF>
   <PowerConsumption />
  </sensitive>
  <sensitive type="k(>=3), l(>=2)">
   <OfficeScale />
  </sensitive>
  <group name="addr" type="quasi" level="k(>=3), l(>=3)"/>
 </head>
 <rdf:RDF>
  <v:Kind>
   <v:fn note="Full Name" />
   <v:bday note="BirthDay" type="quasi" level="k(>=2)" sprit="-" />
   <v:hasTelephone>
    <rdf:Description>
     <rdf:value note="TelephoneNumber" type="open" sprit="\s" />
     <rdf:type note="Number Type" attribute="rdf:resource" number="2" />
    </rdf:Description>
   </v:hasTelephone>
   <v:hasAddress>
    <rdf:Description note="Addresses">
     <v:street-address group="addr" priority="4" />
     <v:locality group="addr" priority="3" />
     <v:postal-code group="addr" priority="2" />
     <v:country-name group="addr" priority="1" />
    </rdf:Description>
   </v:hasAddress>
  </v:Kind>
 </rdf:RDF>
 <OfficeScale note="OfficeScale" />
 <PowerConsumption type="open" note="PowerConsumption" />
</anonymize>
```

Fig. 4 P-XAR Example

publish data samples without data publishing limits to publicize the data's availability. This open information is termed "open attribute." This open attribute can be declared in a data entry.

The secondary data user can request access to the open attributes by utilizing R-XAR. **Fig. 5** lists an example of an R-XAR. If the secondary data consumer requests attributes identified as quasi-identifiers, DAP publishes anonymized data that contains attributes calculated as quasi-identifiers. The user also declares the required anonymization method, privacy protection level, sensitive attributes combinations, open attributes, and quasi-identifiers utilizing the R-XAR.

The formats of XAS and its subset XAR utilize the Cascading Style Sheets (CSS) format and the Semantic Web standard. The XAS can be processed utilizing an XML schema, RDL schema, OWL method, and other related tools.

V. Implementation of PDS

We implemented a DAP application to verify its performance and feasibility. The DAP is the most complicated application, and must be implemented first to enable the evaluation of the proposed infrastructure. The implemented application can confirm the compliance of P-XARs and R-XARs as publishing and requesting rules, respectively. The DAP application utilizes TinyXML-2 [8] for parsing the XAS. P-XARs for the anonymized data were stored in PostgreSQL, a prominent database management system. The application can anonymize data according to the k-anonymity and l-diversity anonymization methods.

Fig. 2 displays a flowchart for the implemented application. Initially, it receives original data as a D-XAS, its publishing rules as a P-XAR, and its requesting rules as an R-XAR. It verifies the format and compliance between the P-XAR and the R-XAR. Subsequently, the application receives the requested data according to the published P-XAR from PostgreSQL, if the data exists. The program then initializes the process of anonymization described in Section II. This program also emulates the process of one-directional anonymization described in Section III. Finally, the program stores the anonymized data as P-XAS into the database.

VI. Evaluation

To evaluate the DAP application, we assumed a typical application example as follows. We utilized the Web access history captured by the designed packet-capturing software implemented on our lab's gateway server. It captures all

```
<?xml version="1.0" encoding="utf-8"?>
<anonymize type="k(3)">
 <head>
  <sensitive>
   <rdf:RDF>
    <v:Kind>
     <v:hasTelephone>
      <rdf:Description>
       <rdf:type number="2" />
      </rdf:Description>
     </v:hasTelephone>
    </v:Kind>
   </rdf:RDF>
   <PowerConsumption />
  </sensitive>
  <group name="addr" type="quasi" />
 </head>
 <rdf:RDF>
  <v:Kind>
   <v:bday />
   <v:hasTelephone>
    <rdf:Description>
     <rdf:value note="TelephoneNumber" type="quasi" />
    </rdf:Description>
   </v:hasTelephone>
  </v:Kind>
 </rdf:RDF>
 <PowerConsumption note="PowerConsumption" />
</anonymize>
```

Fig. 5 R-XAR Example

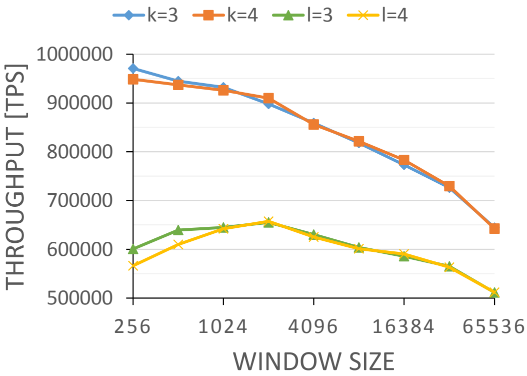


Fig. 6 Throughput of the Implemented Program (1)

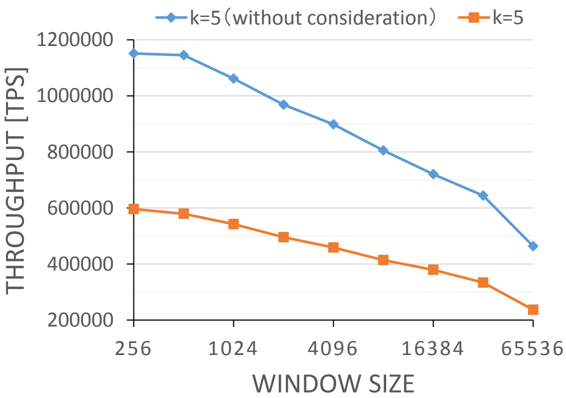


Fig. 7 Throughput of Implement Program (2)

transactions between the Internet and our intranet. Data representing the relationship between the users' IP addresses and the accessed URLs was captured, by referencing the access history. Utilizing this information, we can track the users who frequently access certain URLs. This sample application should manage this captured data as anonymized information. Utilizing this information, we can provide new services, such as a trend survey of Web accesses, a Web page recommendation service, and other useful services.

The application was implemented in C++, utilizing the Debian 7.0 operating system. The host PC includes an Intel Xeon X5650 2.67 GHz CPU, and 36 GB of DDR3 Memory. Test data included two months of Web page access history. The application utilized the first 65,536 instances of all captured accesses, comprising 6,294,273 bytes. In this data, the sensitive attribute is the domain name of Web sites, and quasi-identifier is the user's IP address.

Fig. 6 displays the relationship between the window size and the anonymization throughput of the implemented application. In this example, one-direction anonymization was not utilized; this increases the application's potential vulnerability to privacy breaches. However, the calculation cost is lower, because it calculates k-anonymity or l-diversity only once during the initial anonymizing process. This evaluation measured throughput in tuples per second (TPS). Applying a window size of 4,000, the program can anonymize approximately 600,000 tuples to 800,000 tuples per second. This demonstrates the application is capable of processing large amounts of data.

Fig. 7 additionally displays the relationship between window size and throughput. This diagram, displays the throughput for one-direction anonymization, along with throughput results where one-direction anonymization is not considered. Loss of anonymity was not observed in cases where one-direction anonymization was utilized. However, anonymization throughput was reduced by half when one-direction anonymization was considered; when one-direction anonymization is considered, the program performs the anonymization calculation twice.

VII. Conclusion

We proposed an infrastructure to facilitate the secondary use of data. The proposed infrastructure manages anonymization and the relevant data. We displayed the organizational structure of the proposed anonymization infrastructure and data transactions. This infrastructure prevents further leaks of private information by utilizing one-directional anonymization. We also proposed XAS and its subset, XAR as a format. A protocol to exchange XAS and XAR was also described. The proposed infrastructure will facilitate future services related to the secondary use of data. To evaluate our proposed solution, we implemented an application for DAP. This evaluation demonstrated that the proposed application can anonymize from 600,000 to 800,000 tuples per second, and it exhibits sufficient performance to process large amounts of data.

Acknowledgement

This work was partially supported by Funds for integrated promotion of social system reform and research and development, MEXT, Japan, and by MEXT/JSPS KAKENHI Grant (B) Number 24360230 and 25280033.

References

[1] U.S. Federal Government. DATA.GOV. http://www.data.gov/ (2014.5).

[2] Rakesh Agrawal; Ramakrishnan Srikant, "Privacy-preserving data mining," SIG-MOD, Vol. 29, pp. 439-450, 2000.

[3] Yehuda Lindell; Benny Pinkas, "Privacy Preserving Data Mining," Journal of Cryptology, Vol. 15, pp. 177-206, 2002.

[4] Bee-Chung Chen; Daniel Kifer; Kristen LeFevre; Ashwin Machanavajjhala, "Privacy-Preserving Data Publishing," Foundations and Trends in Databases, Vol. 2, No. 1-2, pp. 1-167, 2009.

[5] Benjamin C. M. Fung; Ke Wang; Rui Chen; Philip S. Yu, "Privacy-preserving data publishing: A survey of recent developments," ACM Computing Surveys (CSUR), Vol. 42, No. 4, 2010.

[6] Latanya Seeney, "k-anonymity: a model for protecting privacy," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 10, No. 5, pp. 557-570, 2002.

[7] Ashwin Machanavajjhala; Daniel Kifer; Johannes Gehrke; Muthuramakrishnan Venkitasubramaniam, "L-diversity: Privacy beyond k-anonymity," ACM Transactions on Knowledge Discovery from Data (TKDD), Vol. 1, No. 1, 2007.

[8] TinyXML-2, http://www.grinninglizard.com/tinyxml2/ (2014.5).

A Low-cost Infrastructure for Massive Storage of Phenotype Data for Dairy Cattle Genetic Improvement Programs

Wagner Arbex*,†, Caio dos Santos Borsato de Carvalho*,‡, Katia Santos*,
Vinicius Campista Brum§ and Marcos Vinícius Barbosa da Silva*
*Brazilian Agricultural Research Corporation (Embrapa)
Juiz de Fora, MG, Brazil
†Correspondent author – wagner.arbex@embrapa.br
‡National Council for Scientific and Technological Development (CNPq) grant
§Federal University of Juiz de Fora (UFJF)

***Abstract*—The activities and progress of genetic breeding were always related to computing – or rather, the availability of appropriate computational resources to the behavior of genetic improvement, because, either by cross breeding or selection, the data sets involved in genetic improvement research and development activities care both in quantity and in quality. The demand for computing resources in these programs is extensive and intense, because genetic-genomic evaluation requires greater availability and accuracy of phenotype data, given the use of assessment models increasingly sophisticated and, moreover, are added to these aspects the need to interpret and understand the phenotype databases from a logical and efficient structure for data storage and information retrieval. This paper presents a low-cost solution to implement a storage system for huge data mass.**

***Index Terms*—Scientific computing environment, storage system, utility computing, phenotype data, genetic improvement**

I. Introduction

Historically, the works and progress of genetic improvement were always related to computing – or rather, the availability of appropriated computational resources to the behavior of genetic improvement [1]. This fact is due to the nature of research activities, because either by cross breeding or selection, the data sets involved in genetic improvement research and development activities care both in quantity and in quality of these.

Addition to the traditional approach of the breeding programs, currently works with genomic selection are being developed, complementing the work already developed and allowing investigation of the genetic potential of individuals before they can express their phenotypic characteristics.

Over again, the demand for computational resources is extensive and intense, because genetic-genomic evaluations require greater availability and accuracy of phenotype data, given the use of assessment models increasingly sophisticated.

Add up to these aspects the need to interpret and understand the phenotype databases from a logical and effective structure for data storage and information retrieval.

Initiatives for data storage in genetic breeding programs are not unknown. Among others, the National Dairy Cattle Research Center (Embrapa Dairy Cattle) of the Brazilian Agricultural Research Corporation (Embrapa) has worked successfully. For example, computing systems were developed for the creation and implementation of the National Animal Science Archive of Dairy Cattle (Arquivo Zootécnico Nacional de Gado de Leite – AZN-GL) estabilished by the Ministry of Agriculture of Brazil through Ordinance #33 on February 10, 1987 [2] .

Now, this work shows a free-software based storage solution to phenotype data warehousing currently being implemented to the Embrapa Dairy Cattle's genetic improvement programs.

II. Data science approach

The computer science and the information systems, utilized as tools by science, changed the agriculture and livestock fields the same way as they did with many other fields of science.

Therefore, considering the exponentially increasing amount and the high complexity of scientific data that are being generated and which need to be efficiently processed, new computational resources are required for effective treatment of this entire data volume, so they can be transformed into knowledge, enabling the adhibition and allowing or enhacing technological advances in order to promote the upgrading of the productive sectors.

The use of mathematical and computational models as a research tool not only makes the interpretation of the handled easily content, but also of complex data sets currently generated, whose characteristics include, inter alia [3]:

- large data volume, which data sets of terabytes magnitude are becoming usual;

- high dimensionality, when working with hundreds or thousands of attributes to be studied;
- heterogeneity, since unlike traditional methods of analysis, computational models are suitable for different data types, discrete and uncategorized;
- multiple physical location of data sets, since it is common these databases are distributed and/or dispersed in different repositories.

III. Motivation and earlier programs

A pioneering initiative to capture phenotype primary data was the development of PROLEITE – a computer software for which three versions were developed [4], [5], [6] – and its use was discontinued in the 90s, partially motivated by the fact that its biggest users, the breeders' associations, began to develop their own applications and/or get them from third-party. These new software could meet their needs and also fulfill the requirement to meet the demand for inclusion of phenotype data in AZN-GL.

The AZN-GL has increased its data volume by 40% per year (Figure 1[1]), at least, and has established itself as a source of data from numerous studies – coinciding with the first studies in Brazil using genetic evaluation software which have implemented the Animal Model.

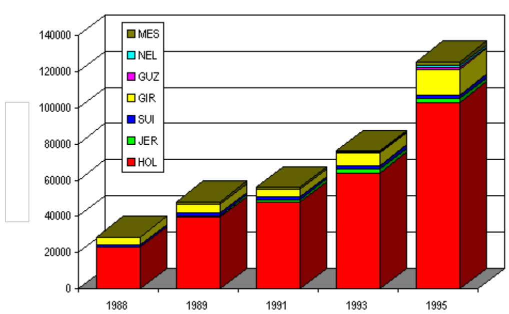


Fig. 1. The AZN-GL growth in the early years of the bovine datasets of the breeds Nelore (NEL), Guzera (GUZ), Gyr (GIR), Brown Suiss (SUI), Jersey (JER), Holstein (HOL) and crossbreed (MES) [7].

In the second half of the 90s, beginning a new venture, the work on the AZN-GL and the PROLEITE became joint projects between the Embrapa Dairy Cattle and the Embrapa Agricultural Informatics [8], [9]. Subsequently, the initiative of the PROLEITE, with new dimensions, was resumed by these research centers [10].

However, the main issue for the maintenance and existence of the AZN-GL was not getting the primary data, but the strategies and methods for the structure and database processing. These strategies and methods should allow the database usage for genetic evaluations implementation – among other analyzes and other research projects investigations.

[1]Nelore, Guzera and Gyr there are Zebu subspecies.

Thus, as another action strategy for the deployment of AZN-GL and genetic evaluations running, a computer system, named Crivo [11], with automatic or semiautomatic methods was developed. These methods were defined in protocols for inclusion, verification, modification, query and retrieval of data.

The methods and strategies deployed in the Crivo proved to be fundamental for all data manipulation of the AZN-GL and, moreover, allowed that their databases were reliable, undamaged and available to be used in several works of the animal breeding research team of Embrapa Dairy Cattle.

Many studies could only be made through the data provided by the AZN-GL. Some of these works are the earliest texts in this field in Brazil, and showed up, e.g., different evaluation models, including the first evaluations with the Animal Model in the country, new studies of reproductive and productive efficiency, the estimation of new parameters for breeds and specific conditions into Brazil, or even the first adjustment factors for lactations also for milk production specific characteristics in the country.

Genetic evaluations using the Animal Model are examples of the importance and necessity of adequate computing resources to process phenotype data. This model, though known since the 60s, became to be used in the late 80s due to the lack of sufficient resources. In Brazil, the Animal Model became to be used in the following decade.

Particularly, new computational resources to perform genetic evaluations using the new Animal Model were required to meet their data preparing needs, and so were added to Crivo new methods that allowed the generation of these files using the new evaluation models template.

It should be emphasized that, at that time, the available computational resources in Brazil were poor – when compared to those used in United States, Canada and European countries that allowed the data processing with some "facility".

In order to, at that time, was necessary to implement non-trivial algorithms that could allow the manipulation of data masses using low computing resources. For example, algorithms for partial and topological sort had been implemented to reduce the memory use, disk access, and thus the runtime.

Algorithms with such implementations were unknown for this purpose – in the case of large data sets – and for small sized hardware. As an example, the development and implementation of such algorithms allowed the work completion conducted by Silva [12] and Houri Neto [13].

The computational activities for Houri Neto's [13] thesis were particularly challenging once datasets with a huge number of entries, up to that time, had not been collected for a single breed in Brazil. In their origin – outside of Brazil – these data sets were hosted in large sized computer equipment, named mainframes, and manipulated by database managers systems (DBMS) commercially established in the market, which was not true in Brazil.

For these reasons, in the past or at the present, the implementation of efficient computational methods that are capable of dealing with this massive amount of data and allow analyze them reliably is characterized as one of the challenges in breeding genetics programs.

IV. Infrastructure for massive storage datasets

Cloud computing is a new business model of IT, but also a new way of computing resources organization where can be made available software, platform, infrastructure, storage area, inter alia. Moreover, it can also provide the desktop computer itself to the end-user. That is, create and provide the desktop resources to users and clients from easy access procedures.

Both technical and operational aspects of cloud computing are based on utility computing, which in turn is structured by a subset of the distributed systems and operating systems concepts and tools, e.g., distribution transparency and virtualization [14], [15].

Storage systems are excellent solutions for scientific computing environments, which usually, the nature of the activity hinders to predict the data volume to be treated. Furthermore, storage systems are expensive solutions and might require changes in the computing environment.

The solution presented was developed using four CPUs with Linux operating system, creating a disk array with capacity of over 4 TB of raw space. To implement the disk array was used MHDDFS [16] and NFS [17] applications, available for Linux and other operating systems. The NFS layer provides the communication among the hard disks over the network and the MHDDFS layer allows the union of several mount points in a single space, assembling the disk array.

Integrated to the storage system was implemented a backup system that copies the whole storage contents, four times a day, in external hard drives with network access. The backup system also had 4 TB of workspace, and it was implemented with native Linux tools – such as CRON and RSYNC – it also has its operation without the intervention or even the user's knowledge.

This storage system was implemented with free software and low-cost hardware, but, despite this, the logical organization and the working structure of the cloud computing allows scalability and security in storage to scientific computing environments and computing on huge databases in general.

Currently, this storage system has been upgraded. It has been improved in the backup mechanism and storage capacity. The storage system area was increased in about 65% and the backup area hard disks was quadruplicate. Respectively, were about 6.5 TB and 16 TB.

Another feature that should be modified is the access way for users, long as new interface and access mechanism will be developed.

V. Conclusions

The efficient and appropriate phenotypic data storage always were problems for genetic improvement programs. Research groups around the world need and are looking for a solution for this issue.

The solution is not restricted to a computational infrastructure for storage resources but begins from this. From an infrastructure point of view this paper presents a low-cost solution using free software and/or open source resources that can be implemented even in obsolete or unsophisticated hardware if necessary. In addition, as a concept of cloud computing and distributed systems, provides storage as a service (StaaS), implementing transparency of access, location, migration, replication, and concurrence.

Possibly, its biggest disadvantage is not to assure a great throughput for data access. Because it does not use an independent or own network structure as many commercial storage systems. However, there are network architectures that can reduce this problem.

The atributes and features offered by this solution can fully meet the needs of massive data storage in genetic breeding programs. Among these, stand out from advantages the ease of implementation, configuration and scalability.

Acknowledgments

The authors thanks to reviewers who gave useful comments, and to funding agencies and corporate sponsors that provided financial support, which will be nominated in the camera-ready version.

References

[1] H. Grosu, L. Schaeffer, P. A. Oltenacu, D. Norman, R. Powell, V. Kremer, G. Banos, R. Mrode, J. Carvalheira, J. Jamrozik, C. Draganescu, and S. Lungu, *History of genetic evaluation methods in dairy cattle*, Jan. 2013.

[2] BRASIL., "Institui o arquivo zootécnico nacional," Diário Oficial, Seção I. Portaria n. 32, de 10 de fevereiro de 1987, Brasília, p. 2126, Fev. 1987.

[3] P. Tan, M. Steinbach, and V. Kumar, *Introdução ao data mining: mineração de dados.* Rio de Janeiro: Ciência Moderna, 2012.

[4] W. Arbex and M. L. Martinez, "PROLEITE/PROLEI-PC 1.0," 2 disquetes, Juiz de Fora, 1989, c e Pascal. Ambiente DOS.

[5] W. Arbex, M. L. Martinez, and M. V. L. dos Santos, "PROLEITE/PROLEI-PC 2.0," 2 disquetes, Juiz de Fora, 1990, c e Pascal. Ambiente DOS.

[6] W. Arbex, M. V. L. dos Santos, and M. L. Martinez, "PROLEITE/PROLEI-PC 3.0," 2 disquetes, Juiz de Fora, 1994, c e Pascal. Ambiente DOS.

[7] S. R. de Medeiros Oliveira, W. Arbex, C. N. Costa, M. A. Arbex, and W. C. P. de Magalhães Júnior, "Automação dos processos de gestão do Arquivo Zootécnico Nacional de Gado de Leite (AZN-GL)," in *Anais...*, Agrosoft 99 – II Congresso da Sociedade Brasileira de Informática Aplicada à Agropecuária e Agroindústria. Campinas: Sociedade Brasileira de Informática Aplicada à Agropecuária e Agroindústria, 1999. [Online]. Available: http://www.agrosoft.org.br/trabalhos/ag99/artigo07.htm

[8] S. R. de Medeiros Oliveira, C. da Costa Carrer, W. Arbex, and J. Valente, "Sistema de análise e acompanhamento de produção de rebanhos leiteiros," in *Anais...*, XXXIV Reunião da Sociedade Brasileira de Zootecnia. Juiz de Fora: Sociedade Brasileira de Zootecnia, Ago. 1997, pp. 338–339.
[9] S. R. de Medeiros Oliveira, M. Pedroso Júnior, W. Arbex, and C. da Costa Carrer, "PROLEITE: sistema de análise e acompanhamento de rebanhos leiteiros," in *Anais...*, Agrosoft 97 – I Congresso da Sociedade Brasileira de Informática Aplicada à Agropecuária e Agroindústria. Belo Horizonte: Sociedade Brasileira de Informática Aplicada à Agropecuária e Agroindústria, 1997, pp. 141–147. [Online]. Available: http://www.agrosoft.org.br/trabalhos/ag97/c3t1100.htm
[10] S. R. de Medeiros Oliveira, C. N. Costa, and C. C. Sabadini, "Sistema para auxiliar o manejo de rebanhos leiteiros e a tomada de decisões," Embrapa Informática Agropecuária, Campinas, Comunicado Técnico 1, Abr. 1999.
[11] W. Arbex and M. L. Martinez, "Crivo," 1 CD, Juiz de Fora, 1996, c. Ambiente Unix-like.
[12] M. V. G. B. da Silva, "Utilização de modelos uni e bivariados no estudo das relações entre eficiência reprodutiva e produção de leite na raças holandesa," Mestrado em Ciências em Zootecnia, Universidade Federal de Minas Gerais, Belo Horizonte, 1995.
[13] M. Houri Neto, "Interação genótipo-ambiente e avaliação genética de reprodutores da raça Holandesa usados no Brasil e nos Estados Unidos da América," Doutorado em Ciência Animal, Universidade Federal de Minas Gerais, Belo Horizonte, 1996.
[14] W. Arbex, R. F. Tagliatti, L. G. de Andrade, M. N. M. Muniz, E. Guedes, and M. V. G. B. da Silva, "Storage as a service and cloud computing for bioinformatics computing environment," in *Anais...*, X-meeting 2010 – VI International Conference of the Brazilian Association for Bioinformatics and Computational Biology. Ouro Preto: Universidade Federal de Ouro Preto - UFO, Out. 2010, p. 103.
[15] W. Arbex, M. V. B. da Silva, M. F. M. Guimarães, R. F. Tagliatti, L. G. de Andrade, M. N. M. Muniz, and L. A. V. de Carvalho, "Storage as a service and utility computing for bioinformatics computing environment: aspects of cloud computing applied to scientific computing," in *Anais...*, IV Encontro Acadêmico em Modelagem Computacional. Petrópolis: Laboratório Nacional de Computação Científica - LNCC, Jan. 2011, trabalho Destaque em Ciência da Computação no IV Encontro Acadêmico em Modelagem Computacional do Laboratório Nacional de Computação Científica.
[16] D. E. Oboukhov, "MHDDFS," 1 file, 2008, MHDDFS – Multiple Hard Disk Distributed File System. [Online]. Available: http://mhddfs.uvw.ru/
[17] S. Shepler, B. Callaghan, D. Robinson, R. Thurlow, Sun Microsystems Inc., C. Beame, Hummingbird Ltd., M. Eisler, D. Noveck, and Network Appliance Inc., "Network File System (NFS) version 4 Protocol," IETF – The Internet Engineering Task Force, Fremont, Tech. Rep. 3530, Dec. 2000. [Online]. Available: http://tools.ietf.org/html/rfc3530

Big Data Anonymization Method for Demand Response Services

Kengo Okada, Hiroaki Nishi

Graduate School of Science and Technology, Keio University, Japan
okada@west.sd.keio.ac.jp, west@sd.keio.ac.jp

Abstract—A demand response services as smart grid application produces and requires large amount of information about electric power consumption. This data can be regarded as big data and is needed to be anonymized for preserving privacy and reducing the amount. Electric power consumption data must be used carefully because it contains private information. The proposed method can convert data to existence probabilities by considering anonymity. In addition, the proposed method can anonymize data according to a required degree of anonymity. An evaluation of demand response services using data anonymized by the proposed method was performed, and the results show that the error rate of the demand and supply balance was smaller than the required level.

Keywords—Anonymization method, Demand response, Privacy Preserving, k-member clustering

I. INTRODUCTION

Recently, owing to the evolution of cloud services, discussion of secondary uses of data has attracted attention, especially for big data. However, preserving privacy is a significant problem. As a solution, an anonymizing technique has been proposed, and the application of this technique to real-world situations has begun. However, for some useful and urgent services, the proposed anonymizing methods do not meet requirements. As a typical application of such services, we have focused on demand response services in a smart grid.

In the area of smart grids, the introduction of a smart meter has been considered and achieved around the world. A smart meter is a new electric power meter, and it is one of the components of home energy management systems (HEMS) [1],[2]. HEMS manage the energy use of a home to achieve a balance between energy saving and a comfortable lifestyle. Smart meters have a communication function to transmit the electric power consumption of a household at regular intervals. The primary use of smart meters is for fare correction. In this case, power consumption data should be collected without loss of data. This means that private information is included in the data.

Demand response (DR) is a typical application of the secondary use of electric power consumption data measured by a smart meter. DR can achieve peak-cut and peak-shift of electricity use by changing the price of electricity or by providing incentives to encourage customers to change their normal consumption pattern when demand for electric power is high. Variable pricing may encourage consumers to reduce electricity consumption and will provide opportunity to think more about how and when we use electricity. Electric companies or aggregators create DR messages to households according to their current electric power demand. An operation test conducted in the United States demonstrated electric power demand reduction of 10–20% [3]. DR can be achieved using electric power consumption data transmitted from residential smart meters.

As an example of DR, electric power consumption will be used widely for secondary use due to its flexibility in applications. The Japanese government has prepared a grant that will introduce more than 10,000 HEMS to a limited city area to encourage secondary use of electric power consumption data and exploit the DR services to facilitate the dissemination of HEMS. The cost of introducing HEMS is comparatively higher than the reduction cost of electricity by itself, which makes it difficult to introduce HEMS widely in Japan.

DR services use a smart meter, which is an electric power meter with a communication function to transmit the electric power consumption of a household to a datacenter. This power consumption data can be used to develop and exploit new services. Recently, secondary use of such data has been considered for such services. DR services are not only for electric companies that collect raw data from smart meters. For other companies, collecting private information with such meters is prohibited. To avoid disclosing private information, it is sufficient for such companies to use generalized or anonymized data if the quality of their services can be guaranteed.

By analyzing household electric power consumption data, companies can provide useful new services. For example, security companies may provide a service that alerts by e-mail when there is no consumption or when consumption is higher than usual when residents are not present at home. In addition, cleaning service of heating, ventilation, and air conditioning appliances can be provided. Such a service can use the air conditioning electric power consumption data to determine clogged filters. Eco-point services, such as discount coupons for various services, can use the data to determine incentives for households that avoid peak use of electricity. Various service providers, such as food service outlets, can also cooperate and share data with electric power companies.

However, it is possible to know what kinds of home appliances are used in a house. Moreover, the family configuration and estimation of income could be analyzed from such data [4]. In a smart grid and clean power conference in Britain, an executive of Siemens Energy said “We, Siemens, have the technology to record energy consumption every minute, second, microsecond, more or less live. From that, we can infer how many people are in the house, what they do, whether they’re upstairs, downstairs, do you have a dog, when do you habitually get up, when did you get up this morning, when do you have a shower: masses of private data.” [5]. If such information is revealed, it may become a threat; e.g., a thief may enter a house when the residents are regularly absent.

Thus, electric power consumption data must be treated with significant care. In this paper, we propose an anonymizing method for electric power consumption data that preserves personal information. The proposed method converts data to distribution data by considering anonymity.

The remainder of this paper is organized as follows. We explain k-anonymization and an existing method for k-anonymization in Section 2. The proposed method is discussed in Section 3, and our evaluation simulation model is described in Section 4. Section 5 presents a hierarchical structure used to treat HEMS data. We present evaluations of the experimental results in Section 6 and finally conclude the paper in Section 7.

II. k-ANONYMIZATION

k-anonymity [6][7] is a generalization method. To discuss k-anonymity, some terms are defined below.

(i) Data table

In this paper, a table-style data expression, such as a database system, is referred to as a “data table.” A column in a data table is an “attribute.” Address, birth, and gender are examples of attributes. One group of data that corresponds to a person or group of people is called a “data set,” and one data set is a “tuple.”

(ii) Attribute

An attribute that can identify a corresponding person, such as name and unique ID, is called an “identifier.” Attributes that cannot be specified on their own, but are identifiable by combining other attributes (e.g., illness, birth, and gender), are referred to as a “quasi-identifier.”

(iii) Sensitive attribute

With the exception of the identifier, an important attribute for secondary use is called the “sensitive attribute.” This attribute is excluded from masking or generalization by anonymization. Moreover, tuple groups that have the same quasi-identifier values are called “q*-block.”

The definition of k-anonymity is as follows. “*Each q*-block in the data table must include at least k tuples.*”

k-anonymity is a privacy protecting model [6][7] that ensures each record of a table is identical to at least $k-1$ other records. For example, we examine two data tables. Table 1 gives patient records. The attributes in this data table are Zipcode, Birthday, Gender, and Disease. Table 2 gives a 2-anonymity (k = 2) version of Table 1. In this table, the sensitive attribute is “Disease”, and quasi-identifiers are “Zipcode”, “Birth” and “Gender”. The data consist of a q*-block from $t1$ to $t3$ and a q*-block of $t4$ and $t5$. This transformation is achieved by generalizing the quasi-identifiers. A quasi-identifier is a subset of attributes that can uniquely identify most tuples in a table.

Even if an attacker wants to know a specific person’s problem and already knows the quasi-identifier of the person, the attacker only can narrow down data to two tuples. As shown in these tables, the masking or generalization process prevents an attacker from identifying a specific person. There are several algorithms for calculating masking or generalization. The most popular algorithm is the heuristic searching method using double nested loops.

The k-anonymization approach can be viewed as a clustering problem. A requirement of k-anonymity can be transformed into a clustering problem so that each cluster contains at least k records. To solve this problem, Byun proposed the k-member clustering algorithm for k-anonymization [8]. First, this algorithm randomly selects a record r as the seed to generate a cluster and adds records to the cluster so that information loss is minimized. Information loss is generally defined as follows.

Let R be a data set and r_i be the ith individual record. Let $e = \{r_1, \dots, r_k\}$ be a cluster where the quasi-identifier consists of numeric attributes $N_1, \dots, N_m$ and categorical attributes $C_1, \dots C_n$. Let $\mathcal{T}_{C_i}$ be a taxonomy tree defined for the domain of categorical attribute C_i. Let Min_{N_i} and Max_{N_i} be the minimum and maximum value in e with respect to attribute N_i, respectively. Let $\cup_{C_i}$ be the union set of values in e with respect to attribute C_i. Then, the amount of information loss incurred by generalizing e, denoted $IL(e)$, is defined as follows:

$$IL(e) = |e|\left(\sum_{i=1,\dots m}\frac{\left(Max_{N_i} - Min_{N_i}\right)}{|N_i|} + \sum_{j=1,\dots,n}\frac{H\left(\Lambda(\cup_{C_j})\right)}{H(\Gamma_j)}\right),$$

Table 1: Patient records

ID	Zipcode	Birthday	Gender	Disease
$t1$	0123	1991.10.29	Female	Cancer
$t2$	0124	1991.6.15	Female	Cold
$t3$	0234	1991.10.24	Male	Flu
$t4$	1110	2004.8.12	Male	Cold
$t5$	1111	2004.6.17	Male	Flu

Table 2: Anonymized records ($k = 2$)

ID	Zipcode	Birthday	Gender	Disease
$t1$	0***	1991	Person	Cancer
$t2$	0***	1991	Person	Cold
$t3$	0***	1991	Person	Flu
$t4$	111*	2004	Male	Cold
$t5$	111*	2004	Male	Flu

where $|e|$ is the number of records in e, $|N|$ represents the size of the numeric domain N, $\Lambda(\cup_{C_j})$ is the sub-tree rooted at the lowest common ancestor of every value in $\cup_{C_j}$, and $H(\mathcal{T})$ is the highest level of the taxonomy tree $\mathcal{T}$.

Once the number of the records in the cluster reaches k, this algorithm selects the furthest record from record r as the new seed and repeats the process to generate the next cluster. When there are fewer than k records that are not registered with any clusters, these records are registered with existing clusters to minimize total information loss. After clustering, these records are generalized in each cluster so that the data table satisfies k-anonymity.

k-anonymity is used to protect privacy data. Electric power consumption data is time-series data. However, an individual can be identified by comparing multiple tables issued continuously even though anonymity is maintained in one table. Therefore, for anonymizing electric power consumption data, k-anonymity is insufficient for preserving individual data. The format of anonymized data is a graph because electric consumption data is originally issued as a graph. If we can obtain anonymized data in a graph, we can use data to create DR messages.

III. Proposed Method

Here, we propose a new anonymizing method for electric power consumption data. This method anonymizes data using the following steps.

1) *k-member clustering*
2) *Convert to existence probability*
3) *Convolution*

Each step is performed as follows:

1) *k-member clustering*

First, the proposed method generates clusters using k-member clustering [8]. If the attributes of records are in a finite numeric domain and the number of the attributes is one, the result of k-member clustering can be described, as illustrated in Figure 1. After k-member clustering, we extract the average and width of each cluster. These data are used in the next step, and the other data are deleted to protect privacy. The results of this step are shown in Table 3. Max_i and Min_i represent the maximum and minimum values of the cluster i, respectively. μ_i is the average of cluster i, and $Width_i$ is width of cluster i.

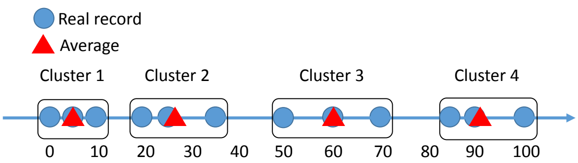


Figure 1: k-member clustering results

Table 3: Results of k-member clustering

Cluster No.	1	2	3	4
Max_i	10	35	70	100
Min_i	0	20	50	85
$Width_i$	10	15	20	15
μ_i	5	26.7	60	91.7

2) *Convert to existence probability*

The proposed method creates existence probability from the average and width of each cluster. We use a Gaussian distribution function to create existence probability. We define existence probability as follows:

$$f_i(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right\}$$

$$\sigma_i = \alpha \cdot Width_i = \alpha(Max_i - Min_i)$$

The existence probability created using the data of cluster i is defined as f_i, and σ_i is calculated using $Width_i$. α is a variable that we set as per a target error value. As α increases, we can obtain a higher degree of anonymity and vice versa. σ represents a dispersion calculated using real data. However, if the dispersion value and the average are used to create existence probability, real data may be leaked by analyzing the dispersion and average. Therefore, we define another σ by using an α value that enlarges the distribution compared with the original dispersion. We use these equations and create existence probability from all clusters. The integral values of the existence probability integral and the Gaussian distribution should both be 1.

For example, we assume two clusters of data with five records each. One cluster includes [40,45,50,55,60], and the

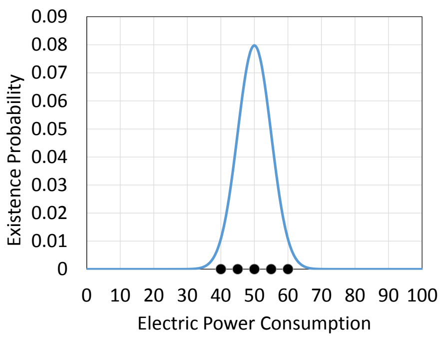


Figure 2: Existence Probability (Sample 1)

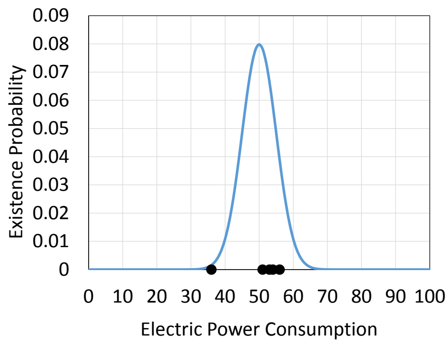


Figure 3: Existence Probability (Sample 2)

other cluster includes [36,51,53,54,56]. In this case, both clusters have an average of 50 and width of 20. Therefore, the existence probability of these clusters has the same distribution, as shown in Figures 2 and 3. These data have anonymity because it is impossible to infer any real data included in each cluster. However, it still remains the tendency of the values included in the cluster because existing probability is created by using average and distribution of each cluster.

3) Convolution

After creating the existence probability for all clusters, a convolution process is required. The proposed method sums all existence probabilities created from each cluster. If cluster number is n, the total area of the existence probability is n because, as described previously, the Gaussian distribution integral value is 1. To adjust total area to 1, we implement scale conversion. This conversion is expressed as follows:

$$F(x) = \frac{1}{n}\sum_{i=1}^{n} f_i(x).$$

$F(x)$ represents the existence probability considering all data and is used as the final anonymized data. The results of convolution are shown in Figure 4.

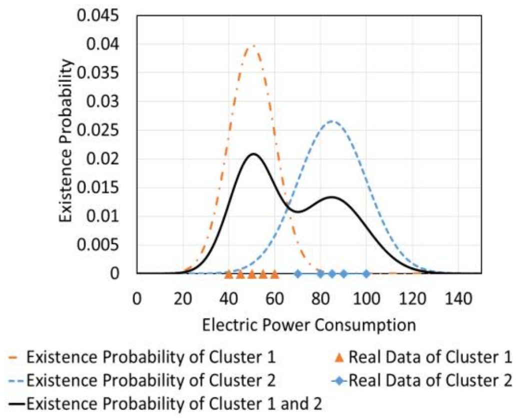


Figure 4: Convolution

IV. SIMULATION MODEL

We evaluate the usability of the data anonymized by the proposed method when used for DR. In this section, we explain the simulation model.

A. Data model

To achieve DR, one major solution is to change the price of electricity or to provide incentives to encourage customers to change their typical consumption pattern when electricity demand is high. To determine the desired amount of reduced electricity consumption, we had to gather all raw electric power consumption data. Using the proposed method, a DR service can be provided without obtaining raw data. Here, we use home electric consumption data and assume a data model similar to that shown in Table 4. By installing smart meters, these data are sent to a server and recorded. We assume the data transmission interval is 30 min because a supply–demand balancing slot of 30 min is employed in Japan. First, the data are transmitted to an electric power company or an aggregator. The companies that use the data for secondary usage obtain the data from the electric power company or the aggregator. However, this causes privacy problems.

Home IDs, such as “home1,” are deleted to avoid specifying individuals when we publish a data table. When generating clusters, it is possible to reveal a home ID. However, this may cause a problem when the cluster is rearranged, and an attacker may distinguish an individual home ID by determining that the specified home ID has been moved to a neighboring cluster.

Table 4: Data Model

time	home1	home2	home3	home4	home5
0:00:00	358.5	218	327	181	470
0:30:00	421.5	209	327	224	368
1:00:00	403.5	206	323	225	382
1:30:00	386	201.5	257	222	294
⋮	⋮	⋮	⋮	⋮	⋮
22:30:00	953	344.5	365.5	947	523.5
23:00:00	516	367.5	263.5	621.5	479
23:30:00	310	324	234	507.5	685

B. Demand response model

In this simulation, the DR service provider issues DR messages to HEMS of all households under the provider to reduce electricity consumption when power demand is high. We assume that all HEMS accept the DR message and reduce electricity consumption automatically after 30 min.

From historical power data trends, we predict electric power demand for the next 30-min interval. To predict this demand, we use both the current data and data from the previous 30-min interval to generate a linear prediction. When the predicted value exceeds a threshold, the system sends a reduction message as a DR message. Note that various studies give more accurate prediction of demand than a linear prediction method. However, the focal point of this simulation is to evaluate the utility of the anonymized data; thus, the accuracy of the demand forecasting method is not essential. Therefore, demand forecasting is performed with linear prediction as a basic prediction method.

C. Creating the reduction command message

Here, the method for creating a DR message is explained. We assume the reduction amount for each household differs by considering fairness, i.e., the rate of household reduction should be higher when the household’s consumption is higher than that of other households.

In this DR simulation, all households are divided into four groups as per power consumption. To conduct this clustering, the existence probability of power consumption data created using the proposed anonymizing method was used. A precise threshold value, distributed to all HEMS of the four household groups, could be clearly determined if real power consumption data is known. However, we can only use obscure anonymized data; thus, the threshold values of the DR groups are calculated as follows.

As mentioned previously, the area of existence probability is 1. For example, if the total number of homes is 50, the responsible ratio of reduction is calculated, as shown in Table 5. Each existence probability is used to determine the power usage threshold by comparing Ratio in Table 5. A responsible Ratio indicates the ratio of the target reduction amount. To divide homes into four groups, we determine three threshold values, as shown in Figure 5. In this case, the threshold values are [530, 390, 250]. The reduction command message is created according to this data.

V. SMART GRID HIERARCHICAL STRUCTURE

The dissemination of HEMS will drastically increase the amount of data available. To seize these data, we propose a hierarchical structure [9]. The proposed anonymization matches this concept well. In Figure 6, data captured in HEMS are transmitted to local nodes. This data includes private information from households in the local area. Usually, the data is forwarded to the master node because this node is generally an Internet backbone router. In this case, the master node must be capable of large calculation and must have significant memory. If the local nodes have a function to gather local data and forward this as anonymized data, the total calculation cost and network congestion rate will be reduced because the amount of data is reduced at the local nodes.

This special local node is implemented as a service-oriented router [10]. The router achieves DPI and has a key-value store-based high-throughput database. It provides an API to deliver services on the router. Using this router as a local node can provide anonymization services within general packet forwarding processes. This special router is provided as middleware of the Juniper MX480 with JunosV App Engine using the Junos SDK Remote API and software model. The HEMS data are converted to existence probability by this router. The calculation cost is not significant because the number of households in the local area is low.

As an evaluation, the master node must collect the data from n households. In the matching process, the calculation cost can be estimated as $O(n^2)$. If it is separated to m local areas, the calculation cost at the master node is reduced to $O((n/m)^2)$. The total calculation cost at a local node becomes $O(nm^2)$. It is clear that $O(n^2) > O((n/m)^2 + nm^2)$.

The proposed anonymization method can also reduce the total data size. For example, we assume that the amount of numeric electric power consumption data in one sample is 8 bytes every 30 min. Then, the amount of data per 30 min becomes 400 Kb when the number of homes is 50,000. On the other hand, the proposed method uses two data in each cluster, i.e., sampled data and anonymized data. If 5-member clustering is used and the number of homes in a cluster is 10,000, the data becomes 160 Kb. Thus, the amount of anonymized data can be reduced by 40% compared to the original data.

When we want to get information of all area, service provides will access the master node and get anonymized data of each area. If we need detail information of each home, the service providers will access local nodes and get detail information in local area. The service providers can access HEMS directly to get plain information by making a special contract with the household. When we use hierarchical structure as shown in Figure 6, we can achieve memory saving and access control. Moreover, efficient services can be provided by adjusting the location to get the data according to the providing services.

Table 5: Ratio of each group

	Number of homes	Rate	Responsible Rate of Reduction (%)
Group1	13	0.26	10
Group2	13	0.26	20
Group3	12	0.24	30
Group4	12	0.24	40
Total	50	1.00	

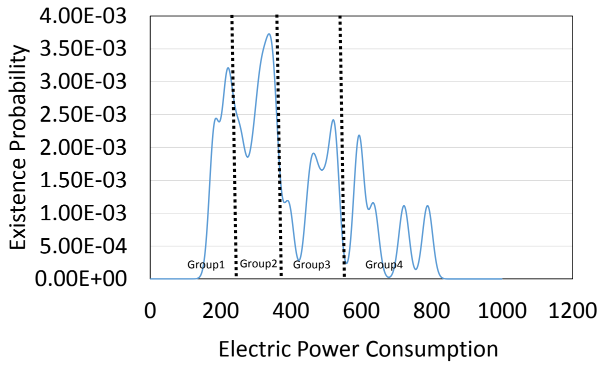


Figure 5: Threshold value for clustering

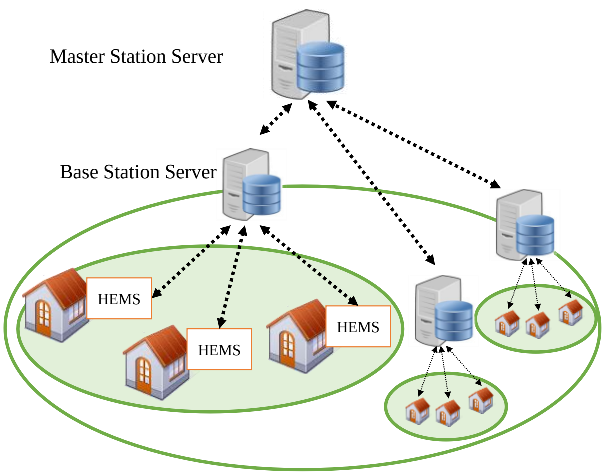


Figure 6: Hierarchical structure in a smart grid

VI. EVALUATION

In this section, the usability of the proposed anonymizing method is evaluated. In the simulation, the Perl programming language was used to handle electric power consumption data, and PostgreSQL was used as a database. These simulations were executed using the CentOS 6.2 Linux operating system. In this evaluation, real electric power consumption data captured from October 2012 to August 2013 was used. This data was generated by HEMS installed in a house in Yukigaya, Tokyo, Japan.

A. Anonymized data

We varied parameters k and α to generate the anonymized data, as shown in Figures 7, 8, and 9. Figure 7 shows the data anonymized by the proposed method for $k = 2$ and $\alpha = 0.5$. In Figures 7, 8, and 9, the dots on the x axis represent real data. The original dots of individuals cannot be identified using the anonymized data.

The value of the existence probability is sufficiently large if part of the distribution is dense and vice versa. As shown in these figures, the anonymized data represent the tendency of real data. The results of anonymizing prove that the degree of anonymity can be increased when parameters α or k are larger, and this can be confirmed by the extension of distribution.

The proposed method secures privacy because it guarantees k-anonymity, and the number of homes is removed.

B. Accuracy of demand response

To evaluate the usability of the data anonymized by the proposed method, we conducted a simulation assuming DR. The error rate of this simulation is defined by the following equation:

$$\text{Error (\%)} = \frac{\sum_{i=1}^{n} |x_{real,i} - x_{anony,i}|}{\sum_{i=1}^{n} x_{real,i}} \times 100$$

Here, n is the number of homes, $x_{real,i}$ is the target value of home i as real data, and $x_{anony,i}$ is the target value of home i as data anonymized by proposed method. When all households observe the required target value, the limitation of total demand was observed.

Figure 10 shows the accuracy of the reduction message in the DR simulation. In this figure, the vertical axis represents k-anonymity, and the horizontal axis represents parameter α, which determines the extent of the existence probability distribution. As the values for parameters k and α increase, the demand response error also increases. For example, when k is 2 and α is 0.5, error becomes 0.61%. On the other hand, when k is 6 and α is 3, error becomes 5.1%. According to the Electricity Business Act of Japan, error of balance between demand and supply of electric power should be less than 3%. Therefore, we set 3% as the target error value in this DR simulation. In Figure 10, the dotted line represents an error rate of 3%. Therefore, if parameters k and α are set under the dotted line, an error that is less than the target error (3%) can be accomplished. Thus, we can anonymize data in consideration of usability.

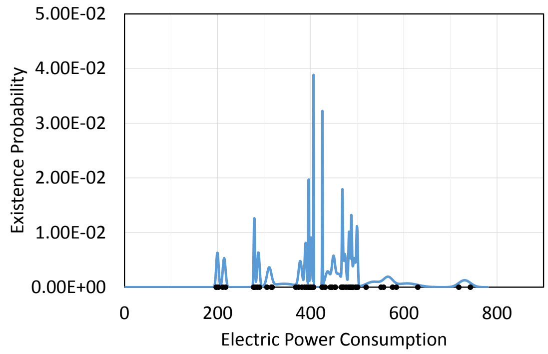


Figure 7: Anonymized data ($k = 2, \alpha = 0.5$)

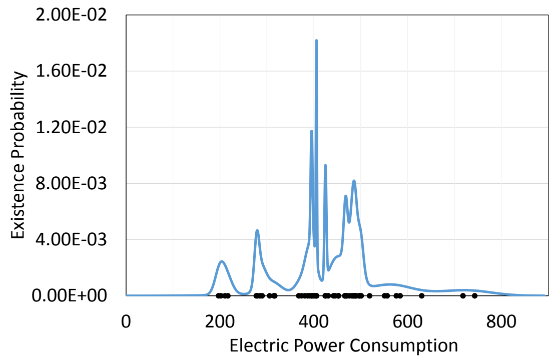


Figure 8: Anonymized data ($\boldsymbol{k = 2, \alpha = 2.0}$)

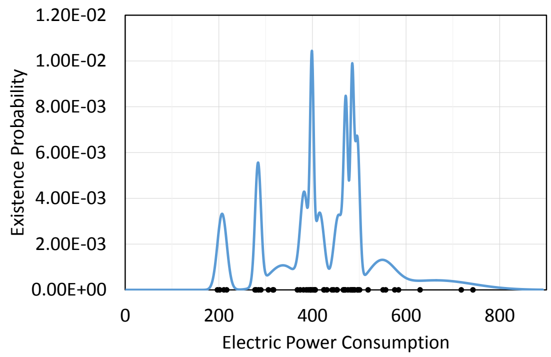


Figure 9: Anonymized data ($\boldsymbol{k = 4, \alpha = 0.5}$)

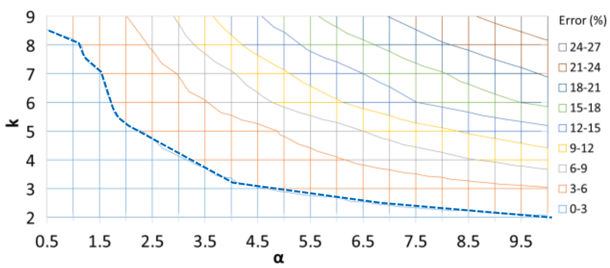


Figure 10: Accuracy of demand response simulation

C. *Amount of anonymized data*

We calculated the amount of data anonymized by the proposed method when the number of homes is 50 and the amount of data is 8 bytes. Table 6 shows the calculation results. In this table, “Amount of Data” represents the amount of anonymized data in a local area per day for 50 homes. “Compression Ratio” represents the result of comparing the amount of anonymized data with the amount of original data. “Compression Ratio is 100%” indicates that the amount of data equals the amount of the original data. As shown in Table 6, as k increases, the amount of data decreases. In other words, as the degree of anonymity increases, the amount of data decreases.

Table 6: Relationship between k and compression ratio (Number of homes: 50)

k	The number of clusters	Amount of Data (KB)	Compression Ratio (%)
2	25	19.2	100
3	16	12.3	64
4	12	9.22	48
5	10	7.68	40
6	8	6.14	32
7	7	5.38	28
8	6	4.61	24
9	5	3.84	20
10	5	3.84	20

VII. CONCLUSION

We have proposed a new anonymization method to preserve the privacy of electric power consumption data. The proposed method converts data to distributed anonymized data as per a dispersion parameter and the average of each group generated using k-member clustering. This k-member clustering satisfies k-anonymity, and the average and width of each cluster are considered to create existence probabilities. These existence probabilities of all clusters are summed and adjusted; thus, the area of existence probability becomes 1.

In a DR simulation, the proposed method could observe a target error value if anonymized data was generated using suitable parameters for k and α. Parameters k and α can vary the degree of anonymity and accuracy of demand control. We demonstrated that a less than 3% error for balancing demand and supply of electric power could be achieved when suitable values for parameters k and α are given. The ranges of parameters k and α that achieve less than 3% error was also demonstrated through an evaluation using real HEMS data.

Secondary use of electric power consumption data is promoted because it can provide useful new services. Companies can provide services that satisfy resident requests for convenience and comfort.

ACKNOWLEDGMENT

This work was partially supported by Funds for integrated promotion of social system reform and research and development, MEXT, Japan,and by MEXT/JSPS KAKENHI Grant (B) Number 24360230 and 25280033.

REFERENCES

[1] M. Kuzlu, M. Pipattanasomporn, S. Rahman, “Hardware demonstration of a home energy management system for demand response applications,” IEEE TRANSACTION ON SMART GRID, vol. 3, no. 4, pp. 1704–1711, DECEMBER 2012.

[2] Y. Ozrurk, S. Datchanamoorthy, S. Kumar, G. Lee, “An intelligent home energy management system to improve demand response,” IEEE TRANSACTIONS ON SMART GRID, vol. 4, no. 2, pp. 1–8, JUNE 2013.

[3] T. Hattori, N. Toda, “Demand response programs for residential customers in the United States—Evaluation of the pilot programs and the issues in practice” MARCH 2011.

[4] “NIST IR 7628 guideline smart grid cyber security vol. 2,” AUGUST 2010, http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf (accessed March 10th 2014)

[5] ELECTRONIC FRONTIER FOUNDATION https://www.eff.org/deeplinks/2010/10/eff-advises-california-puc-smart-grid-privacy (accessed March 10th 2014)

[6] L. Seeney, “k-anonymity: A model for protecting privacy,” International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, vol. 10, no. 5, pp. 557–570, 2002.

[7] K. LeFevere, D. J. DeWitt, R. Ramakrishnan, “Incognito: Efficient full-domain k-anonymity,” Proceedings of the 2005 ACM SIGMOD international conference on Management of data, pp. 49–60, JUNE 2005.

[8] J. -W. Byun, A. kamura, E. Bertino, N. Li, “Efficent k-anonymization using clustering techniques,” Advances in Databases: Concepts, Systems and Applications, pp. 188–200, 2007.

[9] H. Nishi et. al., “IEEE smart grid vision for vehicular technology: 2030 and beyond,” IEEE SMART GRID RESEARCH, IEEE Press, pp. 41–42, 2013.

[10] J. Wijekoon, S. Ishida, E. Harahap, H. Nishi, “Service-oriented router-based CDN system: An SoR-based CDN infrastructure implementation on a real network environment,” IEEE 37th Annual Computer Software and Applications Conference Workshops (COMPSACW), pp. 742–747, JULY 20

Investigation into Indexing XML Data Techniques

Alhadi Klaib, Joan Lu
Department of Informatics
University of Huddersfield
Huddersfield, UK

Abstract- *The rapid development of XML technology improves the WWW, since the XML data has many advantages and has become a common technology for transferring data cross the internet. Therefore, the objective of this research is to investigate and study the XML indexing techniques in terms of their structures. The main goal of this investigation is to identify the main limitations of these techniques and any other open issues. Furthermore, this research considers most common XML indexing techniques and performs a comparison between them. Subsequently, this work makes an argument to find out these limitations. To conclude, the main problem of all the XML indexing techniques is the trade-off between the size and the efficiency of the indexes. So, all the indexes become large in order to perform well, and none of them is suitable for all users' requirements. However, each one of these techniques has some advantages in somehow.*

Keywords: XML Data; indexing XML data techniques; indexing techniques; XML database.

1 Introduction

XML has become a common technology for the transformation of data across the WWW. It was recommended by the World Wide Web Consortium (W3C) in 1998, and has become the standard medium for data presentation and exchange over the WWW. Indexing plays a major role in enhancing XML data queries operation. The relational database is a robust and mature technology and thus more reliable compared with XML. However, the XML data has some advantages compared with the relational model, which are the following: 1) the structure of an XML document is integrated with the data, whereas the structure in a relational model is separate. Thus, it is better to use XML as a medium for transforming data on the Web; 2) XML data has the advantage of flexibility for querying languages, a feature that is not available in SQL; 3) XML data is flexible for adapting to the development of the data structures[1, 2].

1.1 Research questions: this research considers the following questions:

a) What are the main XML data indexing techniques and structures that have been developed and investigated?
b) What are the key results and findings on the XML indexing techniques and structures?
c) What problems, limitations, and challenges that this research area faces?

1.2 Research aims: Broadly speaking, the main goal of this research is to study XML indexing techniques and structures in order to ascertain the key advantages and disadvantages for each of these techniques. In more detail, the following points are the objectives:

a) Investigate the main indexing techniques and schemes.
b) Consider some common comparison criteria for these indexing techniques.
c) Compare these techniques and determine the limitations.
d) Identify the issues and open problems of XML indexing techniques.

1.3 Motivations: The motivation of this research is basically the significance of XML data for database management systems and web technology. Moreover, XML database indexing is a major factor in the efficiency and reliability of XML data technology. This indexing is also a critical demand nowadays due to the growth of XML data usage, XML database size, and the number of users during the last decade. Therefore, it is important to investigate the performance of XML indexing techniques. Thus, this research considers and studies the most popular XML indexing techniques in order to find out the advantages and limitations for each kind of them [3-5].

1.4 Paper organization: The remainder of the paper is structured as follows: Section 2 discusses the research method. The literature review is presented in Section 3. Section 4 discusses the results. Section 5 describes the discussion and the analysis. Section 6 discusses the conclusion, with future work presented in Section 7.

2 Research method

The research method used is a comparison review with a consideration of the research questions, identification of research area, selection process, comparison criteria, and evaluation. The research discusses these stages in the following:

2.1 Review Plan: the first step is to make a plan for the review. This plan considers the method and the process that will be applied. The goal of the study is to review the studied indexing XML data techniques and answer the research questions.

2.2 Identification of Research: a comprehensive and fair search is an important factor to obtain relevant materials such as articles and conference papers, and then to achieve a good review. The start of this search was to identify the search keywords and terms. Different keywords were used in order to obtain as many papers and other materials as possible. The table 1 shows examples of these keywords.

TABLE 1: SEARCH KEYWORDS AND TERMS

XML data	*Indexing XML*
▪ XML principles. ▪ XML data and database. ▪ XML data and semi-structured. ▪ XML Database. ▪ Query XML. ▪ Query optimization. ▪ XML applications. ▪ XML documents.	▪ Indexing XML techniques. ▪ XML indexes classification. ▪ XML Schemes. ▪ XML index structures. ▪ Evaluation of XML indexes.

All possible keywords and terms of the indexing and XML technologies were tried in the search. The Summon (Huddersfield University search engine) and Google Scholar are the most used search engines for the search. The following electronic databases and libraries are the most used resources: ACM Digital Library, Google Scholar database, Science Direct, Springer Link, and Huddersfield University library.

2.3 Selection process: after each search operation, a number of identified articles and papers were ignored since they were irrelevant or duplicate titles. To identify which articles to select, the selection process went through a number of stages as follows:

1- Check the title to find those related to the review.

2- Read the abstract in order to find more details about the articles and exclude all articles not relevant. After conducting this criterion, the number of articles was reduced to about 100.

3- These articles were scanned and considered, as a final check and to exclude any not related ones.

2.4 Quality assessment and classification: the identified papers and articles need to be classified into categories. There are many kinds of classifications for XML indexing techniques, each of which depends on the aims and aspects of the study. The next section gives more detail about the classification and criteria used in this research. With respect to quality, the articles are divided into two classes, namely, development and innovation studies, and analysis and review studies. The criterion for any study being considered as development and innovation study is that the article has proposed or developed a new contribution; whereas the criterion for the analysis and review study is that the article has a survey or a review of others.

2.5 Techniques for classification and review: XML indexing techniques can be classified into three categories according to the ways and aspects in which they are evaluated. First, the indexing techniques are evaluated on the basis of the structural relationships in the index. The classification of this category is usually divided into three classes, as follows: 1) node indexes; 2) path indexes; 3) sequence indexes [6, 7]. Second, this classification is based on the position or location of the index residence. The position can be either in the main memory as a temporary index or in the hard disk, called a disk-based index. The former has the advantage of fast response as it avoids the input and output expenses. However, it has the disadvantage that it lacks scalability for large index files. Third, in this category, classification is based on the type of document that is indexed. There are two classes, namely, the index data-centric document, and the index document-centric XML database [8].

This research focuses on structural relationships indexes as this category is the most relevant to the objectives of the research [9, 10].

2.6 Common criteria for the evaluation of indexing techniques: the ideal method for evaluating a XML indexing technique is to compare it with other techniques using some of the criteria that are suitable for all such techniques. There are a number of common criteria that are used to compare indexing techniques. This research uses some of them in order to evaluate the three structural indexes, namely, node index, graph index and sequence index. These criteria were chosen as the most helpful ones for users to select the best technique for their requirements. The selection is carried out by determining the features that these indexes support, for instance: precision, response time, and completeness. The following are these used criteria:

a) **Retrieval power:** this means the precision and completeness of the result, and the type of queries supported.

b) **Processing complexity**: this step covers a few issues, such as the requirement of structural joins – in order to improve the performance of a query operation; there is a need to minimize the number of joins. Other issues include the processing cost, and the need to compute the relationships between elements.

c) **Scalability**: large indexes involve many input and output operations. Thus this increases the query processing time.

d) **Update cost**: basically, there are two kinds of updates, namely, inserting a node and inserting a subtree. The nodes in a tree index need to be kept organized in a particular way to reflect all kinds of relationships. These relationships have to be preserved if a new node is inserted into the tree. Thus, the index has to reflect its location with respect to these relationships, which makes the case more complex, especially if the scheme has no spaces for the new node.

3 Literature review

XML data is considered to be a semi-structured data since the schema and the data are mixed in the semi-structured data. Thus, this feature provides flexibility for the semi-structured data. Moreover, XML and semi-structured data have more similar features such as simple models, flexibility, self-descriptive models, and readable models for both humans and computers. The semi-structured data is able to represent XML data from other models [1, 11-13].

3.1 XML data indexing techniques: this research focuses on the XML indexes of structural relationships as this category is the most relevant to the objectives of this research. These indexing techniques are evaluated on the basis of the structural relationships in the index. The classification of this category is usually divided into three classes, which are the following: 1) node indexes; 2) path indexes; 3) sequence indexes.

3.1.1 Node indexes: basically, node indexes retain values in the XML tree structure. Each value reflects the location of a node in the tree. These values are used to find a certain node's parents, child, sibling, ancestor and descendent. These values are represented by numbers and used to resolve simple and twig queries. Generally, the most commonly used schemes are the interval (also known as region) labelling scheme and the prefix (also known as path) scheme. Further details about these schemes can be found in the following sections [14-17] [18-21] [22].

3.1.1.1 Prefix scheme: this type of scheme generates code containing two fragments which are the *prefix* part and the *actual-code*. The prefix part encodes the previous node code. The actual-code encodes the order of the node in the tree. There are many examples of this scheme, and the most popular one is *Dewey* labelling. Dewey labelling has two parts in each node except in the root node. This code is called the *Dewey* code [23-25]. The code at each node has two parts. The first part is an increasing number that reflects the location of the node. The second part is the Dewey code which is the parent's code. These parts are separated by a dot like this ".". The root code has only one part since it has no parent. Figure 1 shows an example of a Dewey labelling scheme [16, 24, 26-28].

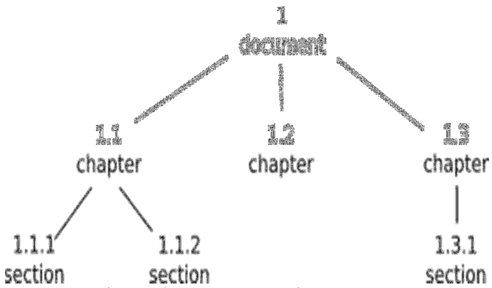


Figure 1. Dewey labelling scheme.

A number of researchers [17] developed a dynamic labelling method that can be used with Dewey labels with identifiers of size 0(log n) where "n" is the size of the database. All labelling schemes including Dewey need 0(n) bits per label [7]. Different kinds of prefix labels are suggested. The following are examples of some of them: Duong and Zhang [29] developed Labelling Scheme for Dynamic XML Data called (LSDX). In this method the numbers and letters are combined. This scheme supports the ancestor/descendent relationship and the sibling between nodes. Lu and Ling [23] propose a labelling scheme that contains two parts. The first part is the group ID. The second part is the group prefix. This scheme is called GRoup base Prefix (GRP). Both the LSDX and GRP schemes are firm and immutable as the label sizes of these schemes can reach 0(n) bits per label. The ORDPATH labelling scheme was developed by a number of researchers [30].

3.1.1.2 Interval labelling Scheme: this is also known as Region-based Encoding. Basically, the idea of this scheme is to attach two values to each node – the startID and the endID. The startID is used to save the node ID for first element or attribute. The endID is used to store the end of the attribute. There are some examples of this labelling scheme such as the (Beg, End) and (Pre, Post) labelling schemes. The (Beg, End) labelling scheme allocates two numbers to each node, based on its sequential traversal order as follows: the mechanism gives a "Beg" number to all elements starting from the root, and each element, attribute of an element, value of an attribute, and value of an element, according to the sequential location in the XML document. When we reach the end of an attribute or an attribute value then the allocated value is the "End" number [31-33].

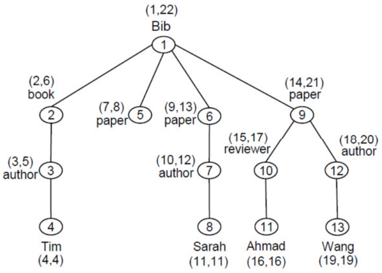


Figure 2. (Beg, End) labelling Scheme [34].

The scheme (Beg, End) can be used to answer both twig query and path query by making use of the relational database management system. This can be achieved by using "structural joins" [35]. To answer a query, the relationships between any couple of nodes in a path in this query is investigated individually as the granularity of this indexing scheme is determined at the level of each node and therefore this provides a precise and complete answer for the query. XML queries shred XML documents into tables of relational databases with the fixed schema (Label, Beg, End, Level, Flag, and Value) [8]. Table 2 represents the relational table of shredded XML document of the above node tree.

TABLE 2: A NODE TABLE OF THE XML DATA IN FIGURE 2 [34]

Label	**Beg**	**End**	**Level**	**Flag (Type)**	**Value**
Book	2	6	2	Element	Null
Author	3	5	3	Value	Tim
Paper	7	8	2	Element	Null
.....	...	...	....	....	....
Paper	14	21	2	Element	Null
Reviewer	15	17	3	Attribute	Ahmad

A number of researchers such as Silberstein et al. and Chen et al. [33] have developed dynamic labelling schemes for interval indexes. The dynamic labelling schemes allow relabeling of the schemes. The interval labelling scheme is the most used scheme for fixed encoding. Such examples propose leaving spaces between the values in order to add new nodes. In case of adding new nodes, there is a need for re-numbering or other solution.

Cohen et al. [7] approved that persistent labelling needs 0(n) bits per label where n is the size of the tree. The interval labels size is used to measure the complexity as this size determines the total size of the index. It is preferable to keep the used number of bits small as this can allow the index to reside in the main memory.

Li and Moon [18] developed the (*Order, Size*) labelling scheme. Each *Order* and *Size* has a certain job. The *Order* one is based on a traversal of pre-order, whereas the *Size* part is an estimation of the number of child or descendent nodes for a given node. The advantage of this mechanism is that this labeling scheme leaves space for any case of adding or inserting nodes in order to avoid relabeling of the data-tree as relabeling can cause delay.

3.1.1.3 Summary: to conclude, both the prefix scheme and interval labelling scheme perform well in XML operations. The interval labelling scheme is better than the prefix scheme in terms of the storage space cost. Thus, some enhancements have been added by reducing the comparison costs and other features in interval labelling scheme. Nevertheless, this scheme is costly in terms of updates [35, 36] [37, 38].

3.1.2 Graph indexing scheme (Path scheme): this type of indexing scheme is also known as a summary index and is described as a structural path summary. It is used to enhance query efficiency by producing a path summary for XML data in order to accelerate the process of query evaluation. However, it can also be used to solve twig queries, but with the extra cost of multiple joins operations. There are a number of graph indexes such as DataGuide [39, 40]; Index Fabric [41]; APEX [42]; D(K)-index [43]; $(F+B)^K$-index [44]; and F&B-index [1, 8]. Graph indexes are classified in different categories according to the number of criteria [41, 42, 45, 46].

Graph indexes have been classified into different types of categories on the basis of different criteria. Examples of these classifications are the following: Polyzotis and Garofalakis classified the graph indexes according to exactness [47]. This classification divided the schemes into exact schemes such as strong Data Guide, 1-index, disk-based F&B-index, Index Fabric, and F&B-index, and approximate schemes. Examples of approximate schemes are A(K)-index, approximate Data Guide, D(K)-index, and $(F+B)^K$-index [47].

There is another category that classifies schemes into two classes, namely, path indexes (aka P-index), which are suitable for simple path queries such as DataGuide and 1-Index, and twig indexes (aka T-index), suitable for twig queries such as F&B-index [8].

Another category [34] considers a classification that categorizes the graph schemes into determinism and bisimilarity. In more detail, in determinism, the paths of the tree are considered to be deterministic paths. The other category, bisimilarity, has two sub classes – forward and backward. According to the determinism and bisimilarity classification, for more detail, graph indexes are categorized into the following classes:

3.1.2.1 Deterministic graph indexes: in this index, every path is listed once in the summary graph. Each path in a summary graph has at least one identical path in the data graph. There are some indexing schemes that are considered as deterministic graph indexes such as Strong DataGuide, Approximate DataGuide, and Index Fabric [39, 40, 43].

The Strong DataGuide was proposed by Goldman and Widon [39]. Strong DataGuides have the ability to give complete and precise results for both simple parent/child path queries and ancestor/descent path. Regarding the twig queries, Strong DataGuides are complete but not precise [44].

The Approximate DataGuide (ADG) has solved the problem of the large size of the Strong Data Guide since this scheme shows large size in some cases [40].

Cooper et al. proposed the Index Fabric in order to solve the problem of scalability [41]. The Index Fabric is theoretically like the Strong Data Guide as the size might enlarge dramatically. Moreover, the Index Fabric is complete for both path and twig queries. However, it is precise for the path but not for the twig [34].

3.1.2.2 Non-deterministic graph indexes with backward bisimilarity: there are a number of these indexing schemes such as: 1-index, A(K) index, and D(K) index. These indexes are divided based on backward bisimilarity. The 1-index was proposed by Milo and Suciu [48] in order to decrease the size of a structural summary. The 1-index divides the data nodes of a document into similar classes based on their backward bisimilarity. The 1-index and DataGuide are the same in the case of s simple XML data tree.

Kaushik et al. [44] proposed the A(K)-index mostly to solve the size cost. The size of an A(K)-index is small compared with Strong DataGuide and 1-index. The A(K)-index is usually complete but not always precise.

Chen, Lim, et al. propose the D(K)-index [43] in order to select the most appropriate value of "*k*" which is a big problem for the A(K)-index. Thus, the D(K)-index is better than the A(K)-index with respect to processing time and storage size. Apart from this, the A(K)-index and D(K)-index are similar schemes in terms of scalability, completeness and precision [34].

3.1.2.3 Non-deterministic graph index with forward and backward bisimilarity: this is the only kind of graph index that has the ability to support twig queries: the F&B-index, $(F+B)^{K-}$index, and the disk based F&B-index [49] [44]. The F&B-index was proposed by Abiteboul et al. [1]. It is different from the A(K)-index and D(K)-index as this scheme is based on the incoming and outgoing paths' bisimilarity for all nodes. Thus, it is a twig structural index scheme. [50] developed the $(F+B)^k$-index. This scheme is an improved release of the F&B-index. The $(F+B)^k$-index scheme controls the size of the F&B-index by identifying the value of the "*K*" [8].

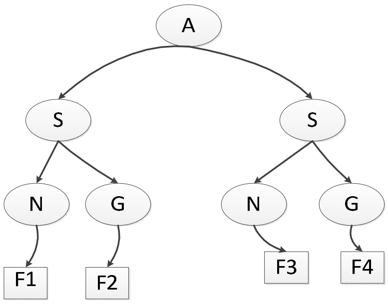
The Disk-based F&B-index was proposed by [51]. This index scheme has provided additional properties and criteria. The Disk-based F&B-index is an integration of 1-index and F&B-index in a new clustered Disk-based F&B-index which is then saved on the disk. [34].

3.1.3 Sequence indexing scheme: this kind of index converts both XML documents and queries into structure sequences. Sequence indexes put the values and the structures of XML data together into an integrated index structure. This new structure is used to evaluate both path and twig queries efficiently: answering a query, making a string sequence that matches the sequence of the data with the query. This technique reduces the need for joins to evaluate twig query [27]. Basically, the sequence schemes are classified into two types according to the importance of tree mapping direction, which are as follows: top-down sequence indexing schemes, and bottom-up sequence indexing schemes.

Wang, Park, Fan and Yu (2003) proposed the ViST. This scheme is based on B+ tree [52]. The ViST (Virtual Suffix Tree) is an example of top-down sequence indexes. The ViST scheme is based on the B+ tree. In addition, the ViST has the disadvantage of weakening the query operations due to the large number of nodes being checked. This is due to the ViST being used as a top-down sequence. Thus, the size of the index becomes very large when dealing with large XML documents since the top elements are added into the sequence. This is the main disadvantage of the ViST scheme. The PRIX (Prufer sequence for indexing

XML) is an example of a bottom-up sequence index. This indexing technique does a good job in decreasing the query processing time. Since the ViST has a problem of scalability as mentioned above, Rao and Moon propose the PRIX as another method that uses a bottom-up sequence to solve the scalability problem with the ViST [53].

Some studies show that these two indexing schemes have a weakness in terms of precision, recall, and processing complexity [52, 54]. However, the sequence indexing techniques have some advantages such as 1) the ability to expect the results of the query; 2) using the complete query tree as one component in order to avoid the cost of the joint operations. Figure 3 is an example illustrates how this indexing technique works.



Label	*Map*
Students	A
PGR	S
Name	N
Group	G
Omar	F1
XDIR	F2
James	F3
SE	F4

Figure 3: Sequence-based indexing examples

From Figure 3, the XML tree is changed into a sequence as follows: T= (A), (S,A), (N,AS), (F1,ASN), (G,AS), (F2,ASG), (S,A), (N,AS), (F3,ASN), (G,AS), (F4,ASG).

Despite the sequence indexing technique having the advantage of speeding the query pre-evaluation and getting rid of the increase of structural joins, this technique also has two disadvantages, namely, 1) the sequence of XML produced by one of these technique algorithms has to be reconstructed quite often; 2) this technique uses a hashing algorithm to encode the textual values, which makes the hash list increase very fast, therefore the indexing becomes very slow [53, 55-60].

4 Results

Using the framework that was defined in the Methods section, the articles were classified accordingly. Table 3 shows this classification and the statistics.

The indexing techniques have been categorized into three classes. Node indexes, graph indexes, and sequence indexes. The investigation in this research has found the following consequences:

1- To the best of our investigation, there is no single indexing technique that is ideal for all users' requirements. Therefore, choosing a suitable indexing technique depends on the user's requirements.

2- With respect to the retrieval power and processing complexity, the node indexing technique is a good one for precision of the results.

3- Regarding scalability, it is clear that most of the indexing techniques suffer from this problem. The problem that these techniques always face is the trade-off between size and efficiency.

4- Update cost is also a common limitation and all investigated techniques have this disadvantage.

TABLE 3: CLASSIFIED ARTICLES

indexing techniques / Articles classifications	Node index	Graph index	Sequence index	Sum	Other articles
Development and Innovation	5	9	7	21	16
Analysis and Review	3	2	2	7	
Sum of each class	8	11	9	28	

4.1 Limitations and open issues of the indexing techniques:

there are number of limitations and open issues that will be discussed in the following:

4.1.1 Limitations: having investigated and studied most indexing techniques, there are three key limitations and problems. These limitations are given as follows:

1- Index size: most of the studied indexing techniques have a problem of scalability. Some indexing techniques are considered to be main memory indexes. Thus, owing to the large size of indexes they cannot reside in the main memory. Furthermore, other approaches that use disk-based indexing techniques also have a problem with the scalability of the indexes as the processing time is affected by the index capacity and its performance.

2- The cost of computation problem: there is a high cost with respect to construction of indexes and the query evaluations procedure. Regarding the relational XML indexes, the disadvantage is that there is a need for a complex computation process in the query evaluation by working out the elements' relationships. The updating of indexes: this problem is a common one among XML indexing techniques.

4.1.2 Open issues and challenges: many researches have been carried out on the indexes of XML data. However, there are still many challenges and open issues that need to be considered. Perhaps the key challenge for XML indexes is the irregularity of structure and data. XML data is considered as semi-structured data. This means that data may be not be complete or may be irregular, and the structure may change quickly and randomly.

5 Discussion and analysis

The XML database systems are a relatively new research area and not as mature as the DBMS. However, many previous studies considered the classifications and evaluations of XML indexing and structures. Each of them has a different investigation and results, depending on the aims and the methodology of the study. Furthermore, XML indexing data is a key factor in enhancing the XML query.

Although heavy research in XML indexing data has been carried out, most of these techniques still face a lack of efficiency. This research has found there is no single technique that is perfect for all types of queries. However, each kind of XML indexing technique has some advantages in different aspects. Node indexes are the most inefficient with regard to structural joins since they need joins for both path and twig queries, whereas graph indexes need no structural joins to support path queries. But for the evaluation of the twig queries, the structural joins are required. Regarding the sequence indexes, they are the most efficient as they encode the structure within the sequence. To summarize, Table 4 shows a summary evaluation for the four techniques using the mentioned criteria.

TABLE 4: COMPARISON BETWEEN THE INDEXING TECHNIQUES USING THE CRITERIA

No.	Criteria	Node index technique	Graph index technique	Sequence index technique
1	**Retrieval power (precision)**	Yes	Yes/no (yes for path & no for Twig)	No
2	**Processing complexity**	No	Yes/no (join required)	Yes
3	**Scalability**	Yes	Yes	Yes
4	**Update cost**	Yes	Yes	Yes

6 Conclusion

To conclude, The XML database systems are not as mature as the relational database management systems which have been studied heavily for decades. The indexing of XML data plays a major role in enhancing the performance of queries. In addition, since XML has a great deal of advantages in terms of data transformation in the WWW, this research has investigated and studied the XML indexing data techniques and structures. The issue is that, to index XML data, it has to represent and reflect the structure, so an efficient path can be made. Moreover, the main problem for indexing techniques is the trade-off between the size and efficiency of the indexes. Therefore, all the indexes become large in order to perform well. The analysis of this research has been carried out on the basis of factors that influence the performance, such as the retrieval power, processing complexity, scalability and update cost. The initial findings show that no one of all the indexing techniques is ideal for all cases and user's requirements.

7 References:

[1] S. Abiteboul, P. Buneman, and D. Suciu, *Data on the Web: from relations to semistructured data and XML*: Morgan Kaufmann, 2000.

[2] R. W. P. Luk, H. V. Leong, T. S. Dillon, A. T. S. Chan, W. B. Croft, and J. Allan, "A survey in indexing and searching XML documents," *Journal of the American Society for Information Science and Technology,* vol. 53, pp. 415-437, 2002.

[3] S. H. Simon, *XML*: McGraw-Hill, 2001.

[4] S. St. Laurent, *XML: a primer*: MIS: Press, 1997.

[5] A. Vakali, B. Catania, and A. Maddalena, "XML data stores: emerging practices," *IEEE Internet Computing,* vol. 9, pp. 62-69, 2005.

[6] N. Bruno, N. Koudas, and D. Srivastava, "Holistic twig joins: optimal XML pattern matching," pp. 310-321.

[7] C. Edith, K. Haim, and M. Tova, "LABELING DYNAMIC XML TREES," *SIAM Journal on Computing,* vol. 39, p. 2048, 2010.

[8] G. Gang, R. Chirkova, and R. Chirkova, "Efficiently Querying Large XML Data Repositories: A Survey," *IEEE Transactions on Knowledge and Data Engineering,* vol. 19, pp. 1381-1403, 2007.

[9] S. Al-Khalifa, H. V. Jagadish, N. Koudas, J. M. Patel, D. Srivastava, and W. Yuqing, "Structural joins: a primitive for efficient XML query pattern matching," pp. 141-152.

[10] M. Atay, A. Chebotko, D. Liu, S. Lu, and F. Fotouhi, "Efficient schema-based XML-to-Relational data mapping," *Information Systems,* vol. 32, pp. 458-476, 2007.

[11] G. Thom, "Data on the Web: From Relations to Semistructured Data and XML," vol. 125, ed: Media Source, 2000, p. 112.

[12] B. Catania, A. Maddalena, and A. Vakali, "XML document indexes: a classification," *IEEE Internet Computing,* vol. 9, pp. 64-71, 2005.

[13] D. DeHaan, D. Toman, M. Consens, and M. Özsu, "A comprehensive XQuery to SQL translation using dynamic interval encoding," pp. 623-634.

[14] M. El-Sayed, K. Dimitrova, and E. Rundensteiner, "Efficiently supporting order in XML query processing," pp. 147-154.

[15] M. Maghaydah and M. A. Orgun, "Labeling XML nodes in RDBMS," pp. 122-126.

[16] T. Härder, M. Haustein, C. Mathis, and M. Wagner, "Node labeling schemes for dynamic XML documents reconsidered," *Data & Knowledge Engineering,* vol. 60, pp. 126-149, 2007.

[17] D. Fisher, F. Lam, W. Shui, and R. Wong, "Efficient ordering for XML data," pp. 350-357.

[18] Q. Li and B. Moon, "Indexing and querying XML data for regular path expressions," in *VLDB*, 2001, pp. 361-370.

[19] M. Al-Badawi, S. North, and B. Eaglestone, "Classifications, Problems Identification and a New Approach," 2007.

[20] M. Al-Badawi, S. North, and B. Eaglestone, "The 3D XML benchmark." vol. 1, ed, 2010, pp. 13-20.

[21] M. Al-Badawi, H. A. Ramadhan, S. North, and B. Eaglestorne, "A performance evaluation of a new bitmap-based XML processing approach over RDBMS," *International Journal of Web Engineering and Technology,* vol. 7, pp. 143-172, 2012.

[22] S. Mohammad, P. Martin, and W. Powley, "Relational universal index structure for evaluating XML twig queries," pp. 116-120.

[23] J. Lu and T. W. Ling, "Labeling and querying dynamic XML trees," in *Advanced Web Technologies and Applications*, ed: Springer, 2004, pp. 180-189.

[24] M. L. Scott and M. L. SCOTT, *Dewey decimal classification*: Libraries Unlimited, 1998.

[25] D. K. Fisher, F. Lam, W. M. Shui, and R. K. Wong, "Dynamic labeling schemes for ordered XML based on type information," in *Proceedings of the 17th Australasian Database Conference-Volume 49*, 2006, pp. 59-68.

[26] I. Tatarinov, S. D. Viglas, K. Beyer, J. Shanmugasundaram, E. Shekita, and C. Zhang, "Storing and querying ordered XML using a relational database system," *Proceedings of the ACM SIGMOD International Conference on Management of Data,* pp. 204-215, 2002.

[27] W. Wei, J. Haifeng, L. Hongjun, and Y. Jeffrey Xu, "PBiTree coding and efficient processing of containment joins," pp. 391-402.

[28] J. Lu, T. W. Ling, C.-Y. Chan, and T. Chen, "From region encoding to extended dewey: On efficient processing of XML twig pattern matching," in *Proceedings of the 31st international conference on Very large data bases*, 2005, pp. 193-204.

[29] M. Duong and Y. Zhang, "LSDX: a new labelling scheme for dynamically updating XML data," in *Proceedings of the 16th Australasian database conference-Volume 39*, 2005, pp. 185-193.

[30] P. O'Neil, E. O'Neil, S. Pal, I. Cseri, G. Schaller, and N. Westbury, "ORDPATHs: insert-friendly XML node labels," pp. 903-908.

[31] P. Dietz, "Maintaining order in a linked list," pp. 122-127.

[32] D. D. Kha, M. Yoshikawa, and S. Uemura, "An XML indexing structure with relative region coordinate," pp. 313-320.

[33] A. Silberstein, H. He, K. Yi, and J. Yang, "BOXes: Efficient maintenance of order-based labeling for dynamic XML data," in *Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on*, 2005, pp. 285-296.

[34] S. Mohammad and P. Martin, "XML structural indexes," Citeseer2009.

[35] C. Zhang, J. Naughton, D. DeWitt, Q. Luo, and G. Lohman, "On supporting containment queries in relational database management systems," *Proceedings of the ACM SIGMOD International Conference on Management of Data,* pp. 425-436, 2001.

[36] X. Wu, M. L. Lee, and W. Hsu, "A prime number labeling scheme for dynamic ordered XML trees," in *Data Engineering, 2004. Proceedings. 20th International Conference on*, 2004, pp. 66-78.

[37] Y. Chen, G. Mihaila, R. Bordawekar, and S. Padmanabhan, "L-Tree: a dynamic labeling structure for ordered XML data," in *Current Trends in Database Technology-EDBT 2004 Workshops*, 2005, pp. 209-218.

[38] B. Yang, M. Fontoura, E. Shekita, S. Rajagopalan, and K. Beyer, "Virtual cursors for XML joins," pp. 523-532.

[39] R. Goldman and J. Widom, "Dataguides: Enabling query formulation and optimization in semistructured databases," 1997.

[40] R. Goldman and J. Widom, "Approximate dataguides," in *Proceedings of the Workshop on Query Processing for Semistructured Data and Non-Standard Data Formats*, 1999, pp. 436-445.

[41] B. F. Cooper, N. Sample, M. J. Franklin, G. R. Hjaltason, and M. Shadmon, "A fast index for semistructured data," in *VLDB*, 2001, pp. 341-350.

[42] C.-W. Chung, J.-K. Min, and K. Shim, "APEX: An adaptive path index for XML data," in *Proceedings of the 2002 ACM SIGMOD international conference on Management of data*, 2002, pp. 121-132.

[43] Q. Chen, A. Lim, and K. W. Ong, "D (k)-index: An adaptive structural summary for graph-structured data," in *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, 2003, pp. 134-144.

[44] R. Kaushik, P. Bohannon, J. Naughton, and H. Korth, "Covering indexes for branching path queries," pp. 133-144.

[45] "XRel: a path-based approach to storage and retrieval of XML documents using relational databases," *ACM Transactions on Internet Technology (TOIT),* vol. 1, pp. 110-141, 2001.

[46] M. Yoshikawa, T. Amagasa, T. Shimura, and S. Uemura, "XRel: a path-based approach to storage and retrieval of XML documents using relational databases," *ACM Transactions on Internet Technology,* vol. 1, pp. 110-141, 2001.

[47] N. Polyzotis and M. Garofalakis, "Structure and value synopses for XML data graphs," in *Proceedings of the 28th international conference on Very Large Data Bases*, 2002, pp. 466-477.

[48] T. Milo and D. Suciu, "Index structures for path expressions," in *Database Theory—ICDT'99*, ed: Springer, 1999, pp. 277-295.

[49] H. Su-Cheng and L. Chien-Sing, "Evolution of Structural Path Indexing Techniques in XML Databases: A Survey and Open Discussion," pp. 2054-2059.

[50] R. Kaushik, P. Shenoy, P. Bohannon, and E. Gudes, "Exploiting local similarity for indexing paths in graph-structured data," in *Data Engineering, 2002. Proceedings. 18th International Conference on*, 2002, pp. 129-140.

[51] W. Wang, H. Jiang, H. Wang, X. Lin, H. Lu, and J. Li, "Efficient processing of XML path queries using the disk-based F&B index," in *Proceedings of the 31st international conference on Very large data bases*, 2005, pp. 145-156.

[52] H. Wang, S. Park, W. Fan, and P. Yu, "ViST: a dynamic index method for querying XML data by tree structures," pp. 110-121.

[53] P. Rao and B. Moon, "PRIX: Indexing and querying XML using prufer sequences," in *Data Engineering, 2004. Proceedings. 20th International Conference on*, 2004, pp. 288-299.

[54] H. Wang and X. Meng, "On the sequencing of tree structures for XML indexing," in *Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on*, 2005, pp. 372-383.

[55] S. Mohammad, "INDEX STRUCTURES FOR XML DATABASES," Dissertation/Thesis, 2011.

[56] P. Rao and B. Moon, "Sequencing XML data and query twigs for fast pattern matching," *ACM Transactions on Database Systems (TODS),* vol. 31, pp. 299-345, 2006.

[57] J.-M. Bremer and M. Gertz, "Integrating document and data retrieval based on XML," *The VLDB Journal,* vol. 15, pp. 53-83, 2006.

[58] X. Meng, Y. Jiang, Y. Chen, and H. Wang, "XSeq: an indexing infrastructure for tree pattern queries," in *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, 2004, pp. 941-942.

[59] K. H. Prasad and P. S. Kumar, "Efficient indexing and querying of XML data using modified Prüfer sequences," in *Proceedings of the 14th ACM international conference on Information and knowledge management*, 2005, pp. 397-404.

[60] B. Stein, "Principles of hash-based text retrieval," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 2007, pp. 527-534.

SESSION

MOBILE COMPUTING AND APPLICATIONS

Chair(s)

TBA

Effective Nutrition Label Use on Smartphones

Vladimir Kulyukin
Department of Computer Science
Utah State University
Logan, UT, USA
vladimir.kulyukin@usu.edu

Tanwir Zaman
Department of Computer Science
Utah State University
Logan, UT, USA
tanwir.zaman@aggiemail.usu.edu

Sarat Kiran Andhavarapu
Department of Computer Science
Utah State University
Logan, UT, USA
sarat.kiran@aggiemail.usu.edu

***Abstract*—Proactive nutrition management is considered by many nutritionists and dieticians as a key factor in reducing and controlling cancer, diabetes, and other illnesses related to or caused by mismanaged diets. As more and more individuals manage their daily activities with smartphones, smartphones have the potential to become proactive diet management tools. Many grocery products sold worldwide have nutrition labels (NLs). Unfortunately, even highly motivated consumers sometimes find it difficult to locate or to comprehend them. The literature on NL use by consumers contains several recommendations to improve retention and comprehension of nutrition information: 1) central positions of NLs; 2) nutrients sorted by health relevance; 3) explanation of nutrients; 4) reduced visual clutter around NLs; 5) increased visual salience through contrast and orientation; 6) increased surface size of NLs. In this paper, a system is presented that satisfies recommendations 1, 3, 4, and 6. The system's front end is implemented as a smartphone application. The smartphone application runs on the Google Nexus 4 smartphone with Android 4.3 or 4.4. The system's back end is currently a four node Linux cluster used for image recognition and data storage. The presented system has broader implications for food policy. The position advocated in this paper argues that the current NL design on product packages does not necessarily have to change to make NL use more effective. Rather, consumers can use their smartphones to design and manipulate their own NL presentation schemes suitable to their specific nutrition needs without requiring product manufacturers to change physical product packages.**

***Keywords*—mobile computing; cloud computing; nutrition label use; nutrition management; electronic commerce**

I. Introduction

U.S. Department of Agriculture estimates that U.S. residents have increased their caloric intake by 523 calories per day since 1970 [1]. Mismanaged diets are estimated to account for 30-35 percent of cancer cases. A leading cause of mortality in men is prostate cancer. A leading cause of mortality in women is breast cancer. Approximately 47,000,000 U.S. residents have metabolic syndrome and diabetes. Diabetes in adults and children appears to be closely related to increasing obesity levels. It is estimated that by 2030 the prevalence of diabetes in the world will reach 4.4%, which will equal to approximately 366 million people [2]. Due to the long-term complications of diabetes, many countries will likely see an increase in blindness, kidney failures, and amputations. Many nutritionists and dieticians consider proactive nutrition management to be a key factor in reducing and controlling cancer, diabetes, and other illnesses related to or caused by mismanaged diets.

Many products sold worldwide have nutrition labels (NLs). In the U.S., the display of nutrition information is mandated by the Nutrition Education and Labeling Act (NLEA) of 1990 [3]. Similar initiatives or legislative acts (e.g., EU FLABEL [4]) exist in other countries. Unfortunately, even highly motivated consumers, who look for NLs to make healthy food choices, sometimes find it difficult to locate and to comprehend nutrition information on many products [5]. Recent investigations of NL use by consumers have used digital cameras to track consumers' eye movements to better understand how consumers locate and understand NLs [6]. These studies have identified four key factors that appear to impede comprehension and retention of nutrition information: 1) label's location on the package; 2) presentation of information within the label; 3) label's surface size; and 4) surrounding visual clutter. Consumers report that they can better locate NLs positioned centrally on a side with a small amount of surrounding visual clutter. Consumers also report failures to comprehend nutrition terms and to read small font sizes in NLs [7].

Several recommendations are made in the NL use literature to improve retention and comprehension of nutrition information: 1) central positions of NLs; 2) nutrients sorted by health relevance; 3) explanation of nutrients; 4) reduced visual clutter around NLs; 5) increased visual salience through contrast and orientation; 6) increased surface size of NLs [5].

In this paper, a system is presented that satisfies recommendations 1, 3, 4, and 6. The system's front end is implemented as a smartphone application. The application runs on the Google Nexus 4 smartphone with Android 4.3 or 4.4. The system's back end is currently a four node Linux cluster used for image recognition and data storage.

The front end smartphone sends captured frames to the back end cluster across a wireless data channel (e.g., 3G/4G/Wi-Fi) where barcodes, both skewed and aligned, are recognized [10]. Corresponding NLs are retrieved from a cloud database, where they are stored as HTML documents, and sent across the wireless data channel back to the smartphone where the HTML documents are displayed on the touchscreen. Wikipedia links to important nutrition terms are embedded for better comprehension. Consumers can use standard touch gestures (e.g., zoom in/out, swipe) available on mainstream smartphone platforms to manipulate the label's surface size. The NL database currently includes 235,000 products compiled from public web sites by a custom crawler.

The remainder of this paper is organized as follows. In Section II, the system's overview is presented and vision-based nutrition extraction methods are discussed to give the reader a broader background about the front end of the system. In Section III, the node cluster is described in detail. Section IV presents several stress test experiments with the system and discusses the results. Section V presents our conclusions, outlines the strengths and limitations of our system, and discusses several implications for proactive nutrition management and food policy.

II. Related Work

A. Overview

Modern nutrition management systems assume that users understand how to collect nutritional data and can be triggered into data collection with digital prompts (e.g., email or SMS) [7]. Such systems often underperform, because many users find it difficult to integrate nutrition data collection into their daily activities due to lack of time, motivation, or training. Consequently, they eventually turn off or ignore numerous digital stimuli [8].

To overcome these challenges, we have begun to develop a Persuasive NUTrition Management System (PNUTS). PNUTS seeks to shift current research and clinical practices in nutrition management toward persuasion, automated nutritional information extraction and processing, and context-sensitive nutrition decision support.

PNUTS is based on a nutrition management approach inspired by the Fogg Behavior Model (FBM) [8], which states that motivation alone may not be insufficient to stimulate target behaviors such as nutrition intake recording or blood tests. Even a motivated user must have both the ability to execute a behavior and a trigger to engage in that behavior at an appropriate place and time. Many nutrition management system designers assume that consumers and patients are either more skilled than they actually are or that they can be trained to obtain the required skills. Since training is difficult and time consuming, a more promising path is to make target behaviors easier and more intuitive to execute.

PNUTS makes proactive nutrition management easier and more intuitive by utilizing the relative advantages of mobile and cloud computing to improve nutrition information comprehension and retention and to automate real-time vision-based NL analysis and nutrition intake recording [9, 10]. In this paper, we focus on effective NL use on smartphones that addresses four out of six major factors that impede nutrition information retention and comprehension by consumers. While vision-based nutrition intake recording is beyond the scope of this paper, we give a sketchy overview of how it works to give the reader a broader background on the system.

B. Barcode Scanning

In PNUTS, there are two kinds of nutrition information extraction algorithms: barcode localization and scanning and nutrition information extraction from NLs. In this section, we give a sketchy overview of both algorithms. Interested readers are referred to [9] and [10] for technical details and experiments.

Figure 1. Skewed barcode

Figure 2. Localized skewed barcode

Recognized barcodes are used to retrieve NLs from a database of NLs. In our previous research [11], we presented an eyes-free algorithm for vision-based localization and decoding of aligned barcodes. The algorithm was based on the assumption that simple and efficient vision techniques, when augmented with interactive user interfaces, ensure that the smartphone camera is horizontally or vertically aligned with the surface on which a barcode is sought.

In [10], we presented two algorithms that relaxed the horizontal or vertical alignment constraints, which may not always hold, to localize skewed barcodes in frames captured by the smartphone's camera, as shown in Figures 1 and 2. The first algorithm localizes skewed barcodes in captured frames by computing dominant orientations of gradients (DOGs) of image segments and collecting smaller segments with similar DOGs into larger connected components.

The second algorithm localizes skewed barcodes by growing edge alignment trees (EATs) on binary images with detected edges. Since our experiments showed that the DOG algorithm outperformed the EAT algorithm [9], the current version of PNUTS uses the DOG algorithm for barcode localization. The localized barcodes are scanned without any image rotation. Our current barcode scanning algorithm handles both UPC and EAN formats. Unlike other barcode scanning solutions (e.g., https://github.com/zxing/zxing, http://redlaser.com), our algorithm does not require the user to align the smartphone's camera with the barcode and can detect skewed or aligned barcodes anywhere in the image.

C. Nutrition Label Segmentation

We have also been working on a vision-based algorithm to localize NLs on grocery product packages and to segment them into text chunks for subsequent optical character recognition (OCR) [9]. The algorithm captures frames in video mode from the smartphone's camera, localizes horizontally or vertically aligned NLs (see Figure 3), and segments the NLs into single- or multi-line text chunks, as shown in Figure 4 (right). Each text chunk is given to an OCR engine. While we have been using free open source GOCR (jocr.sourceforge.net) and Tesseract (http://code.google.com/p/tesseract-ocr/) engines, other OCR engines can be used as well.

Figure 3. Horizontally and vertically aligned NLs

Images captured from the smartphone's video stream can be divided into foreground and background pixels. In general, foreground pixels are defined as content-bearing units in a domain-dependent manner. For example, content can be defined as black pixels, white pixels, pixels with specific luminosity levels, specific neighborhood connection patterns (e.g., 4-connected, 8-connetected), etc. Background pixels are those that are not foreground.

The horizontal projection of an image (HP) is a sequence of foreground pixel counts for each image row. The vertical projection of an image (VP) is a sequence of foreground pixel counts for each column in an image. Figure 4 shows the vertical projection of an NL image after edge detection, which is done in our system with the Canny Edge detector [11].

In detecting NL boundaries, three assumptions are currently made: 1) an NL is present in the image; 2) the NL present in the image is not cropped; and 3) the NL is horizontally or vertically aligned. The detection of NL boundaries proceeds in three stages. Firstly, the first approximation of the vertical table boundaries is computed. Secondly, the vertical boundaries computed in the first stage are extended to the left and to the right. Thirdly, the upper and lower horizontal boundaries are computed.

The objective of the first stage is to detect the approximate location of the NL along the horizontal axis. This approximation starts with the detection of horizontal lines in the image, which is accomplished with a horizontal line detection kernel (HLDK) that we developed in our previous research and described in our previous publications [12, 13]. It should be noted that other line detection techniques (e.g., Hough transform [14]) can be used for this purpose. Our HLDK is designed to detect large horizontal lines in images to maximize computational efficiency.

Let HLFI be a horizontally line filtered image, i.e., the image put through the HLDK filter or some other line detection filter (see Figure 4 left). The projections of white pixels can then be computed for each column of HLFI. The right image in Figure 4 shows the vertical projection of the HLFI on the left. A threshold is chosen, which in our application is set to the mean count of the white foreground pixels in columns. In Figure 4 (right), the threshold is shown by a red line. It can be observed that the foreground pixel counts in the columns of the image region with the NL are greater than the threshold.

The vertical boundaries of an NL are computed as follows. Firstly, the left boundary is extended to the first column to the left of the current left boundary, for which the projection is at or above the threshold, whereas the right boundary is extended to the first column to the right of the current right boundary, for which the vertical projection is at or above the threshold.

A typical NL includes text chunks with various types of caloric and ingredient information, e.g., "Total Fat 2g 3%." To optimize the performance of subsequent OCR, which is beyond the scope of this paper, these text chunks are segmented from localized NLs (see Figure 4). This approach is flexible in that segmented text chunks can be wirelessly transmitted to multiple cloud servers for parallel OCR. As can be seen in Figure 5 (left), text chunks are separated by black colored separators. Formally, text chunks are defined as image segments separated by horizontal black separator lines. The chunking algorithm detects such separators as well as areas with high concentration of corners between the separators, the theory being that text segments are areas with higher concentrations of corners. Figure 5 (right) shows text chunks detected in the localized NL shown in Figure 5 (left). These chunks are subsequently sent to an OCR engine [13].

D. Nutrition Label Crawler

Our original intention was to use only computer vision to extract NLs to populate the NL database. Both the localization and text chunking procedures outlined in the previous section showed robust performance levels in our experiments [13]. Unfortunately, when these techniques were coupled with open source OCR engines such as GOCR (jocr.sourceforge.net) and Tesseract (http://code.google.com/p/tesseract-ocr/), the number of OCR errors was high.

The crawler module was implemented to compensate for low OCR rates. We hope that, as OCR rates improve, the need for the crawler component will become less significant, because NL databases will be populated with computer vision. Another reason that we are currently working to improve OCR rates is that web site scraping may be unreliable in the long term in that a site that currently permits robots may prohibit them in the future. Additionally, some sites may contain inaccurate or obsolete nutrition information.

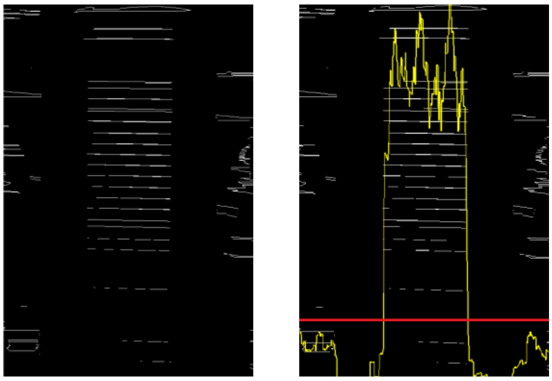
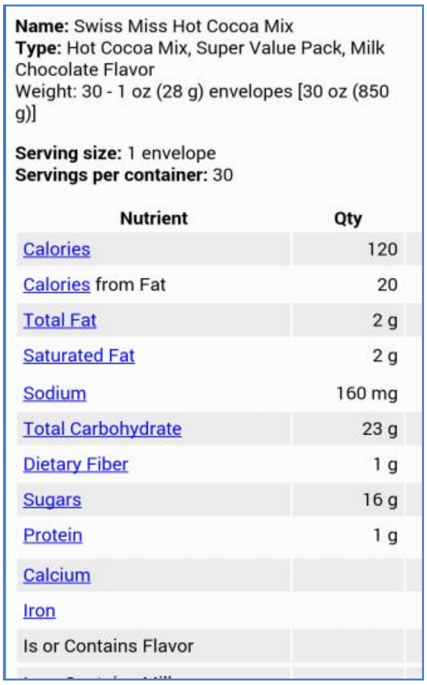


Figure 4. Edges detected in NL (left) and its vertical projection (right)

Figure 5. Segmented NL (left) and its text chunks (right)

The current version of the crawler scrapes three public web sites dedicated to nutrition, which we found helpful in our background research on public nutrition information sites: www.directionsforme.org, www.smithsfoodanddrug.com, www.digit-eyes.com. As we find other helpful public sites, we may add them to the crawler's list in the future. Permissions are verified for each web site before it is scraped.

Each URL is parsed with the Python BeautifulSoup library (http://www.crummy.com/software/BeautifulSoup/) which is implemented on top of the popular Python XML and HTML parsers LXML (lxml.de) and HTML5LIB (http://code.google.com/p/html5lib/) . Each URL is parsed to obtain the NL, ingredients, warnings, and categories. When this information is extracted, a new HTML document is generated that contains not only the extracted information but also embedded Wikipedia links to all nutrition terms used in the tabular part of the NL. A path to this HTML document is saved in a database under a specific barcode. The NL database currently includes 200,000 products compiled from public web sites by the crawler.

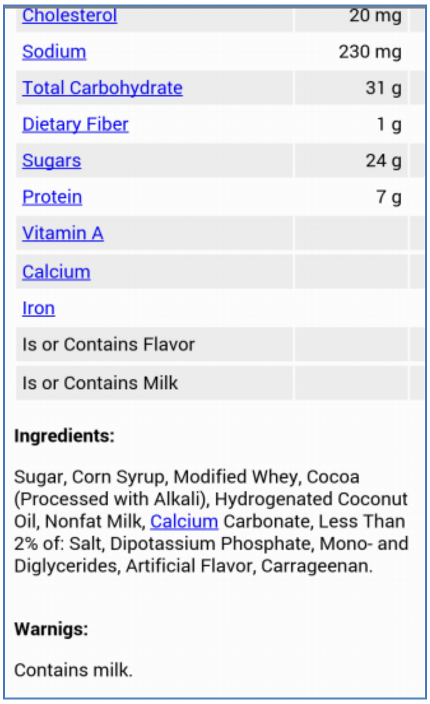


Name: Swiss Miss Hot Cocoa Mix
Type: Hot Cocoa Mix, Super Value Pack, Milk Chocolate Flavor
Weight: 30 - 1 oz (28 g) envelopes [30 oz (850 g)]

Serving size: 1 envelope
Servings per container: 30

Nutrient	Qty
Calories	120
Calories from Fat	20
Total Fat	2 g
Saturated Fat	2 g
Sodium	160 mg
Total Carbohydrate	23 g
Dietary Fiber	1 g
Sugars	16 g
Protein	1 g
Calcium	
Iron	
Is or Contains Flavor	

Figure 6. Upper part of NL with embedded Wiki links



Cholesterol	20 mg
Sodium	230 mg
Total Carbohydrate	31 g
Dietary Fiber	1 g
Sugars	24 g
Protein	7 g
Vitamin A	
Calcium	
Iron	
Is or Contains Flavor	
Is or Contains Milk	

Ingredients:

Sugar, Corn Syrup, Modified Whey, Cocoa (Processed with Alkali), Hydrogenated Coconut Oil, Nonfat Milk, Calcium Carbonate, Less Than 2% of: Salt, Dipotassium Phosphate, Mono- and Diglycerides, Artificial Flavor, Carrageenan.

Warnigs:

Contains milk.

Figure 7. Lower part of NL with embedded Wiki links

E. Nutrition Label Display

When a barcode is recognized at the back end in an image sent to it from the smartphone, its HTML document (if there is one) is sent back to the smartphone and displayed on the touchscreen, as shown in Figure 6.

Consumers can use standard touch gestures (e.g., zoom in/out, swipe) for manipulating the label's surface size or browsing its contents. For example, Figure 7 shows the lower part of the NL displayed in Figure 6 after the user does a down swipe on the touchscreen. When the user clicks on an embedded link, a Wiki page for that nutrient is displayed, as shown in Figure 8.

This presentation method satisfies four out of the six recommendations made in the NL use literature to improve nutrition information retention and comprehension. As stated in Section I, there are six recommendations: 1) central positions of NLs; 2) nutrients sorted by health relevance; 3) explanation of nutrients; 4) reduced visual clutter around NLs; 5) increased visual salience through contrast and orientation; 6) increased surface size of NLs.

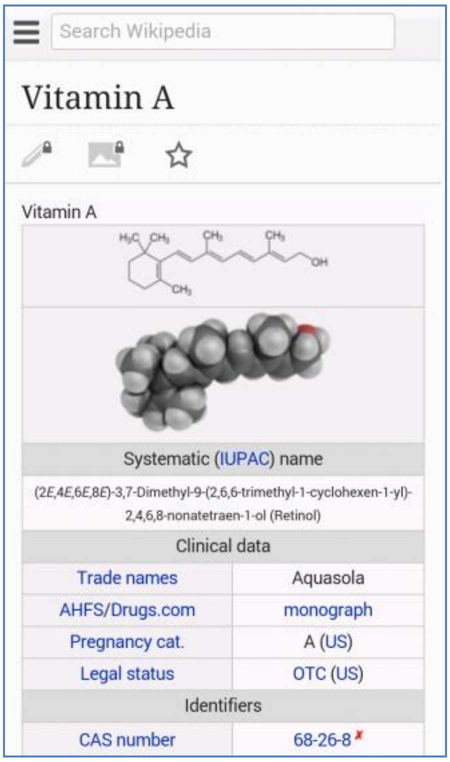


Figure 8. Wiki page of a nutrient (Vitamin A)

This presentation addresses the first recommendation by positioning NLs centrally on the touchscreen, as shown in Figures 6 and 7. The third recommendation is addressed through embedded Wiki links to the nutrients in the tabular component of each NL, as shown in Figure 8. The fourth recommendation is completely satisfied, because there is no visual clutter on the touchscreen around the displayed NL. The sixth recommendation is also addressed, because the user can use standard touch gestures to increase or decrease the actual size of the NL, which can be beneficial not only for regular users but also for low vision ones.

III. Linux Node Cluster

To implement the back end of our system, we have built a Linux cluster out of four Dell computers for cloud-based computer vision and data storage. Each computer has an Intel Core i5-650 3.2 GHz dual-core processor that supports 64-bit computing. The processors have 3MB of cache memory. The machines are equipped with 6GB DDR3 SDRAM and have Intel integrated GMA 4500 Dynamic Video Memory Technology 5.0. All machines have 320 GB of hard disk space. Ubuntu 12.04 LTS was installed on each machine.

We used JBoss (http://www.jboss.org) to build and configure the cluster and the Apache mod_cluster module (http://www.jboss.org/mod_cluster) to configure the cluster for load balancing. Our cluster has one master node and three slaves. The master node is the domain controller. The master node also runs mod_cluster and httpd. All four machines are part of a local area network and have hi-speed Internet connectivity. We have installed JDK 7 in each node.

The JBoss Application Server (JBoss AS) is a free open-source Java EE-based application server. In addition to providing a full implementation of a Java application server, it also implements the Java EE part of Java. The JBoss AS is maintained by jboss.org, a community that provides free support for the server. JBoss is licensed under the GNU Lesser General Public License (LGPL).

The Apache mod_cluster module is an httpd-based load balancer. The module is implemented with httpd as a set of modules for httpd with mod_proxy enabled. This module uses a communication channel to send requests from httpd to a set of designated application server nodes. An additional communication channel is established between the server nodes and httpd. The nodes use the additional channel to transmit server-side load balance factors and lifecycle events back to httpd via a custom set of HTTP methods collectively referred to as the Mod-Cluster Management Protocol (MCMP).

The mod_cluster module provides dynamic configuration of httpd workers. The proxy's configuration is on the application servers. The application server sends lifecycle events to the proxies, which enables the proxies to auto-configure themselves. The mod_cluster module provides accurate load metrics, because the load balance factors are calculated by the application servers, not the proxies.

All nodes in our cluster run JBoss AS 7. Jboss AS 7.1.1 is the version of the application server installed on the cluster. Apache httpd runs on the master with the mod_cluster-1.2.0 module enabled. The Jboss AS 7.1.1 on the master and the slaves are discovered by httpd.

A Java servlet for image recognition is deployed on the master node as a web archive file. The servlet's URL is hardcoded in every front end smartphone. The servlet receives images uploaded with http post requests, recognizes barcodes, and sends the appropriate HTML pages back to front end smartphones. The HTML files generated by the crawler are stored on the cloud via a shared directory implemented as a Network File System (NFS) on the cluster. No data caching is currently done on the servlet or the front end smartphones.

IV. Experiments and Results

We tested the robustness of the node cluster in a series of stress test experiments. The objective was to check the accuracy of our cluster configuration and load balancing. We took eight Google Nexus 4 smartphones from our laboratory running Android 4.3 or 4.4 and deployed a node cluster stress tester application on each of them.

The application would start a background service at startup. The background service would download a random 1024 x 1024 barcode image from an http server, upload it to the node cluster, and display the node cluster's response at the smartphone's action bar.

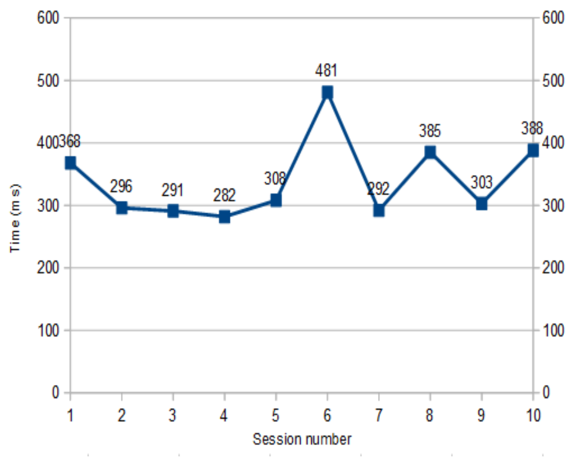


Figure 9. Node cluster request-response pairs

The applications on both smartphones were executed for ten sessions of 3,000 request-response pairs each and the average request-response time for each session was calculated. Figure 9 gives the graph of the node cluster's request-response times. The lowest average was 282 milliseconds; the highest average was 481 milliseconds.

Additional node cluster testing was done in an undergraduate mobile application development class taught by the first author at Utah State University in the fall 2013 semester. One of the assignments asked the students to write an image uploader application to stress test the node cluster. Specifically, fourteen students, each of whom had an Android smartphone, implemented and deployed this application on their smartphones and ran it for one week. Each application would submit three images per second for a total of forty images per second. Thus, during this week, the node cluster was tested with eight lab smartphones and fourteen student smartphones and received a total of sixty six 1024 x 1024 images per second (twenty four images from the lab smartphones and forty two images from the students' smartphones). The node cluster did not experience any failures and was able to handle and balance the load.

V. Conclusions

The R&D literature on NL use by consumers contains several recommendations for improving nutrition information retention and comprehension: 1) central positions of NLs; 2) sorting of nutrients by health relevance; 3) explanation of nutrients; 4) reduced visual clutter around NLs; 5) increased visual salience through contrast and orientation; and 6) increased surface size of NLs.

In this paper, we presented how our system, called PNUTS, addresses recommendations 1, 3, 4, and 6 to increase the effectiveness of NL use on smartphones. The system leverages vision-based barcode recognition to retrieve NLs for specific barcodes. Wikipedia links to important nutrition terms are embedded in NLs positioned centrally on the smartphone's touch screen. The user can follow the links to improve comprehension and retention of nutrition information. The system leverages the standard touch gestures (e.g., zoom in/out, swipe) to enable the user to manipulate the label's surface size and browse NLs. The NL database currently includes 230,000 products compiled from public web sites by a custom crawler.

PNUTS currently does not address recommendations 2 and 5. To address recommendation 2, user profiles will have to be added to the system. For example, if a user has Type II diabetes, the system can automatically sort the nutrients in each NL according to some relevancy taxonomy worked out in collaboration with a dietician. Alternatively, a smartphone UI can be designed to enable the user to specify the health relevance of nutrients for subsequent display.

A similar approach may turn out successful in addressing recommendation 5. For example, visual salience of displayed NLs can be increased by coding important NL components (e.g., carbohydrates, dietary fiber, sugar, etc.) with different colors or display them in a pie chart. It should also be possible to enable the user to choose a visual salience enhancement pattern at configuration time.

The average request-response time between the front end and the back end was three seconds. We expect additional reductions in request-response times to come from faster data communication plans (e.g., 4G) and adding additional nodes to the cluster. Additional time reductions will come from data caching both on the front end and the back end. In the current version of the system, no data caching is currently done on the node cluster or the smartphones. The smartphones can, for example, maintain a local cache of barcodes and retrieved NLs and display NLs without receiving them from the node cluster.

The presented system has an important implication for proactive nutrition management and food processing industry. A major implication for proactive nutrition management is that PNUTS enhances the user's ability to record and comprehend nutritional intake. The user is no longer required to manually enter either names or barcodes of consumed products. In the future, we plan to extract caloric information from NL HTML files or, in the longer term, from captured images automatically.

Our system also has implications for broader food policy. The recommendations for improving NL use appear to focus on product manufacturers. The central theme of these recommendations appears to be that the product manufacturers should be rationally or legislatively persuaded to change the NL design on product packages. However, in order to change the NL design on a physical package, the manufacturer must bear many costs such as disruptions in product recognition and, quite possibly, reduced advertisement space on the package. Moreover, even if product manufacturers adopt a different NL design, there is no guarantee that the one-size-fits-all approach will succeed with all consumers. There will always be consumers who may not like the new design and may prefer additional or different design customization.

The approach presented and advocated in this paper argues that the current NL design on product packages does not necessarily have to change to make NL use more effective. Rather, the strengths of mobile and cloud computing can be leveraged to increase the effectiveness of NL use. Consumers can use their smartphones to design their own NL presentation schemes suitable to their specific nutrition needs without requiring product manufacturers to change physical product packages. While it remains to be seen which of these two approaches will be more successful in the long run, the wide adoption of smartphones and cloud services by the public at large is an indicator that the approach presented in this paper has potential.

References

[1] World Health Organization. "Annual World Health Statistics." Avail. at http://www.who.int/gho/publications/world_health_statistics/en/.

[2] Rubin, A. L. Diabetes for Dummies. 3rd Edition, Wiley, Publishing, Inc. Hoboken, New Jersey, 2008.

[3] Nutrition Labeling and Education Action of 1990. Avail. at http://en.wikipedia.org/wiki/Nutrition_Labeling_and_Education_Act_of_1990.

[4] Food Labelling to Advance Better Education for Life. Avail. at www.flabel.org/en.

[5] Graham, D. J., Orquin, J. L., and Visshers, V. H. M. "Eye tracking and nutritional label use: a review of the literature and recommendations for label enhancement," Food Policy, vol. 32, pp. 378-382, 2012.

[6] Graham, D.J. and Jeffery, R.W. "Location, location, location: eye tracking evidence that consumers preferentially view prominently positioned nutrition information," J. Am. Diet. Assoc., vol. 111, pp. 1704–1711, 2011.

[7] Årsand, E., Tatara, N., Østengen, G., and Hartvigsen, G. 2010. Mobile Phone-Based Self-Management Tools for Type 2 Diabetes: The Few Touch Application. *Journal of Diabetes Science and Technology*, 4, 2 (March 2010), pp. 328-336.

[8] B. J. Fog. "A behavior model for persuasive design," In Proc. 4th International Conference on Persuasive Technology, Article 40, ACM, New York, USA, 2009.

[9] Kulyukin, V., Kutiyanawala, A., Zaman, T., and Clyde, S. "Vision-based localization and text chunking of nutrition fact tables on android smartphones," In Proc. International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV 2013), pp. 314-320, ISBN 1-60132-252-6, CSREA Press, Las Vegas, NV, USA, 2013.

[10] Kulyukin, V. and Zaman, T. "Vision-based localization of skewed upc barcodes on smartphones," In Proc. International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV 2013), pp. 344-350, pp. 314-320, ISBN 1-60132-252-6, CSREA Press, Las Vegas, NV, USA, 2013.

[11] Canny, J.F. "A Computational approach to edge detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, 1986, pp. 679-698.

[12] Kulyukin, V., Kutiyanawala, A., and Zaman, T. "Eyes-free barcode detection on smartphones with niblack's binarization and support vector machines," In Proc. 16-th International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV 2012), vol. I, pp. 284-290, CSREA Press, July 16-19, 2012, Las Vegas, Nevada, USA, (pdf); ISBN: 1-60132-223-2, 1-60132-224-0Y, 2012.

[13] Kulyukin, V., Vanka, A., and Wang, W. "Skip trie matching: a greedy algorithm for real-time ocr error correction on smartphones," International Journal of Digital Information and Wireless Communication (IJDIWC), vol 3(3), pp. 56-65, ISSN: 2225-658X, 2013.

[14] Duda, R. O. and Hart, P. E. "Use of the hough transformation to detect lines and curves in pictures," Comm. ACM, vol. 15, pp. 11–15, January, 1972.

An investigation into the problems of user oriented interfaces in mobile applications

A. Ragab Ihnissi [1] **and B. Joan Lu**[2]
[1] School of Computing and Engineering, University of Huddersfield,
Huddersfield, West Yorkshire, UK

Abstract – *The purpose of this paper is the analysis and evaluation of the mobile interface design. This study consisted of a random sample of 55 user interfaces for mobile applications. In addition, the restriction of all the components of the user interface quantified. An analysis was conducted of these interfaces, in order to represent graphically. Then, evaluated and produced the following results: First, the smaller number of pages in the application is better. Second, decreasing the navigation bars, buttons and menus in user interfaces for mobile applications gives additional space on the screen, making the application easy to use and maintaining the context. Third, diversity, the use of tools ensures good interaction with the user. Finally, a range of results for the design of the user interface and some ideas are provided about what should be taken of these results in mind when designing interfaces for mobile applications.*

Keywords: User interface, mobile application

1 Introduction

Mobile phones have grown to be overwhelmingly popular within today's modern society. Rising sales involving mobiles for that reason, has turned into an important goal for several corporations[1]. A lot of people have more than one mobile appliance such as smart phone, tablet and PDA [2]. The recent changes in the environment are to learn an important requirement in the event. Mobile phones have become popular in the community, many people can tolerate the cost[3].

However, a wireless network and computer network technologies has been developed, learn from face-to-face, distance learning, development evolved and mobile learning. Therefore, mobile learning is unparalleled in that it provides real and personalized learning everywhere, at any time, and so provides teachers and students a chance to get through the simple operation of any and all class materials on their mobile devices[4]. The particular raising quantity in addition to increasing functionality of such gadgets has generated pattern problems because of the modest sizing, insufficient one on one selection potential, in addition to moderate pattern standardization[5]. On the other hand, conventional user interface knowing is not adequate to develop efficient interfaces for cell phone applications, because the mobility context presents developers with many new challenges and peculiarities. [6]. Even so, quite a few usability complications have been earned by the multi-functionality regarding mobile phones.

There was an important raise for the UI usability and design. While using advantages regarding many different varieties and keypads inside cell phones, this user's need to have regarding controllability, grip-stability, and usability have been lifted from the user interface aspect[7]. User interfaces are still suffering of some problems including:

- Lack of making use of screen space.
- Small keyboard buttons result in errors when the user enters data.
- Congestion data reduces user interaction.
- Some user interfaces with an intricate design, making it lose its appeal to users.
- Lack of use voice and video in most applications.
- The problems introduced inside the habits collection will try to cover both the most common problems when making mobile UIs nowadays, as well as long term challenges similar to multi-modal (including utilization of gestures) in addition to contextual as well as adaptive UIs[8, 9].

This article aims to analyse the mobile phone interface design and evaluation. It debates many significant design problems, and describes emerging technologies to promote the effective design and development methods and user-friendly interface for mobile applications appear. Authors collected a random sample of 55 mobile user interface applications for analysis and evaluation.

2 Related work

Nilsson structured that user interface design is proposed models of a collection of mobile applications. Also proposed the models collection of solutions to a number of problems may arise when design such solutions[8] . Bertelsen and Nielsen stated that defies phone models and user interface, classification and augmented reality interface technology used as a "thinking tool", the development of ideas and interaction for mobile devices[10]. Gong and Tarasewich argued that features and limitations of the current mobile apparatus interfaces, particularly as contrasted to a desktop environment[11].

Park et al suggested that specified proceedings and methods to assist manner directory the development of mobile phone user interface guidelines[12]. Jin and Ji proposed that availability of methods of risk assessment and the physical user interface used in this study can help designers identify and evaluate design features and physical user interface of critical materials[7]. Uzunboylu et al stated that the use of mobile technology and multimedia information systems combine and data services, to raise the development of environmentally and mobile technology conscious[13]. Park et al Investigated two pilot studies to improve the emotional connection to our understanding of the movement of the reaction, the quality aspects of the mobile user interface touch screens[14].

Browne and Anand evaluated that three user interface used to play an iPod Touch scrolling shooter video game has been an experimental test of efficiency and capability to enjoy[15]. Morris and Tomlinson stated that the development of A graphical user interface, which is to improve the working mobile phones, other portable devices and personal digital assistants[5]. Liarokapis and Conradi discovered that portable navigation effective in the urban environment and find the path, when explore the possibility of using the user interface lightweight[16]. Lumsden and Global stated that focus on the different methods of user interface adaptation of equipment, with particular attention to enable architectures to adapt to the user's predilections and the environments[17]. Huber et al proposed that the use of the design space for mobile video browsing broaden horizons[18]. Motiwalla evaluated the fusion distance learning or traditional classroom environments of mobile technology[19]. Huang et al studied that provides Mobile Plant Learning System (MPLS) for teachers avenues and means to promote student learning cycle in botany at the primary level[20].

3 Results and analysis

In this paper, these were collected from a random sample consisting of 55 interfaces and mobile applications, where it was limited to the components of each user interface separately quantified. In addition, values were sorted on the basis of how much application use this tool from the sample as a whole. On the other hand, the calculated percentage looks like this: the number of applications that have used this tool divided by the total number of samples multiplied by 100. Accordingly, it has been the representation of these results in graphical forms for the purpose of analysis and evaluation.

Figure 1 displays the classification model components and user interface, which has been divided into three main parts: first, Process Control includes navigation bar, buttons, menu and toll bar. Secondly, contents included page, table, chart and text box. Thirdly, vision included text view, image view and video view.

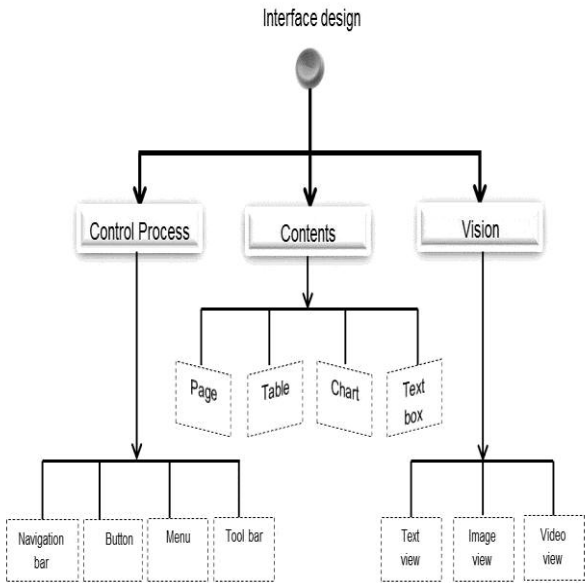


Fig 1. Model of user interface design[21]

$$\sum_1^n \mathrm{I}[\mathrm{P} \subseteq, \mathrm{C} \subseteq, \mathrm{V} \subseteq] \qquad (1)$$

$$\mathrm{P} \subseteq \{\mathrm{p}_1 \ldots \mathrm{p}_\mathrm{n}\}, \mathrm{C} \subseteq \{\mathrm{c}_1 \ldots \mathrm{c}_\mathrm{n}\}, \mathrm{V} \subseteq \{\mathrm{v}_1 \ldots \mathrm{v}_\mathrm{n}\}$$

I: interface, P: process, C: contents, V: vision

Navigation bar: varies from one application to another and according to the nature of the application and thinking designer. Sample reveals that 11 interfaces are not included navigation bar on their pages like Gigabyte's GSmart[22] and there are others in most of its pages, such as Smart pay[23].

Button: it seems clear from sample that one application does not contain buttons is TriplAgent [54], and another application that contains the largest number of buttons is Camera Genius 4.2[24] containing 63 buttons. For all other applications, the number of buttons was anisotropic due to the different purposes of these applications, such as Email.me.it[25] contains 4 buttons, TrustGo[26] contains 18 buttons, Viber[27] contains 30 buttons.

Menu: sample illustrates that 19 of the applications are not containing lists such as Smart pay, Samsung and SelfieCam[23, 28, 29]. This means that the multiplication of the menus within one application makes it complex and difficult to use and it becomes unattractive to the user. On the other hand, some applications contain 9 lists such as FlatPlayer[30], this shows that the application contains many options. In addition, the high number of menus in the application makes the user lose context so that it does not interact well with him.

Tool bar: sample displays that 19 of the applications only used this tool, as it gives more flexibility to the user and reduces congestion buttons on the page. In addition, enables the user access to plug-ins for the application.

Page: the sample shows that all interfaces contain at least one page like Dribbble shot[31] and a maximum of 12 pages such as Időkép[32]. This does not mean that the application specified a certain number of pages, which determines the number of pages is the type of application and coverage of all the aspects that set-up for it, taking into account ease of use and clarity. But the smaller number of pages creates a better user interface and enables the user to maintain context and interaction with the application.

Table: the sample reveals that the tables are almost non-existent in most applications, 3 applications only contain tables of 55 applications such Email.me.it, Smart pay and MIX[23, 25, 33]. This explains that most of the applications in this sample do not use tables to display data from the data source (database), but use other tools such as a text box.

Chart: note from the sample that 13 of the applications of the total sample used charts such as Tide, Reportly, Analytics and statistics[34-36]. Limited use of graphs of statistical data and represent it graphically.

Text box: this tool was used in 14 of the applications that you need to enter data, view or using a user name and password it also describes the sample, such as Email.me.it, Smart pay and Messenger[23, 25, 37].

Text view: the sample offers that the majority of applications contain a text view and some of them reached 17 like Viber, Quartier Senegalais and Taxt[27, 38, 39]. In addition, 8 applications are not contain text view such as Dribbble shot, FlatPlayer and Samsung[28, 30, 31]. Usefulness of text view displayed some of the details of the task or function illustrates a specific tool.

Image view: the sample shows that 20 of the applications did not use the image view because these applications display information, numbers or graphs. However, the rest of the other applications exposed images either as a means to illustrate, for the announcement or advertising applications such as cafe, restaurant or any other product like The porter beer bar, BurgerQuest[40, 41]. In addition, used widely in applications of image editor such as Camera Genius 4.2[24].

Video view: this tool is used with applications that display TV channels and video clips or sports applications, it also showed the sample that four applications only, use this tool of the total sample: a Discovery Channel, FlatPlayer, citizen.tv and Liga Moche[30, 42-44].

Figure 2 shows the number of pages used in applications. A closer look at the data reveals that applications using three pages had the highest percentage of approximately 22 per cent. In addition, followed by applications using a one-page reached 16%, and four pages of 14.5%. On the other hand, were lower rates for applications that used the 5.6, 8, 9, 10, 11 and 12 pages ranged between 2% and 5.5%. However, there were applications used two pages about 11%, while almost 9% of the applications used 7 pages.

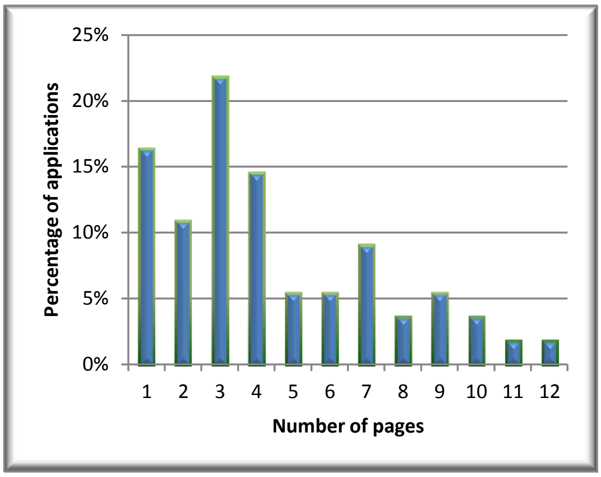


Fig 2. The number of pages for applications

Figure 3 illustrates the percentage of the number of times using the navigation bar in applications. The highest percentage was for applications that do not use navigation bar reached 20%, the second-highest percentage of applications that are used navigation bar once or twice reached 16%. 14% of the applications used navigation bar four times, and 13% used it three times. Some applications used five times by 5%. 4% for each of the applications are used 7,8 and 11 times, while the ratio is at least 2% of the applications that are used 6 and 9 times.

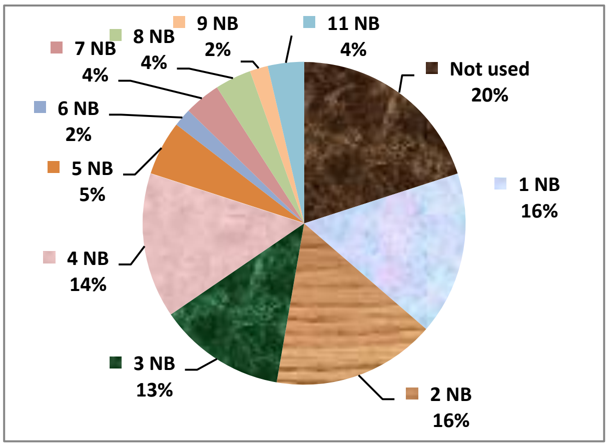


Fig 3. Percentage to use the navigation bar in applications

From Figure 4 that the values of the buttons of varying and capricious, and starts from only one application does not contain any button such TriplAgent[45], while another application has the highest value of the buttons and reached the 63 buttons such as Camera Genius 4.2[24]. Also note that most of the applications contained 2-21 buttons like Tidean and Half Centric[34, 46]. At the same time, eight applications included 22-37 buttons such as Free Ringtones and Liga Moche[44, 47].

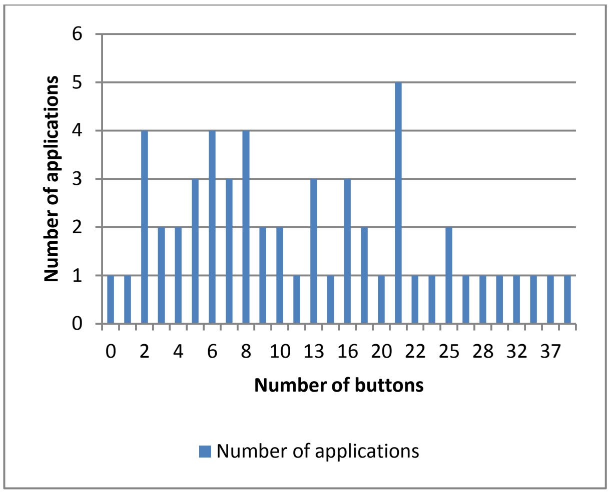


Fig 4. Number of buttons for applications

The graph provided reveals the percentage to use menus in UI for mobile applications [Fig 5].It is noted that most of the user interfaces for mobile applications do not use menus by 34%, while 27% reached for user interfaces using a single list. Followed by applications that use three menus of 13%. On the other hand, there are a small percentage of applications that use a large number of menus, where it was 9% of the applications used four lists and 4% used five menus, while the rest of the applications that used 6, 7 and 9 menus were 2%.

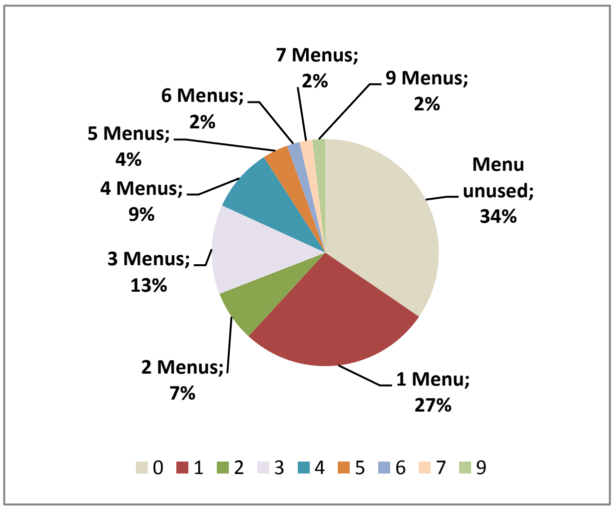


Fig 5. The Percentage to use menus in applications

A glance at the graph provided reveals the percentage to use some tools in UI for mobile applications [Fig 6]. Most applications have used the text view tool where the largest percentage about 85.5%, followed the image view tool almost 64%. On the other hand, a smaller percentage of the table tool was 5.5%, followed by a video view tool 7%. The applications used a chart tool were modestly compared with the rest of the tools almost 24%, and 27% of the text box tool. 35% of the applications have used the toolbar tool.

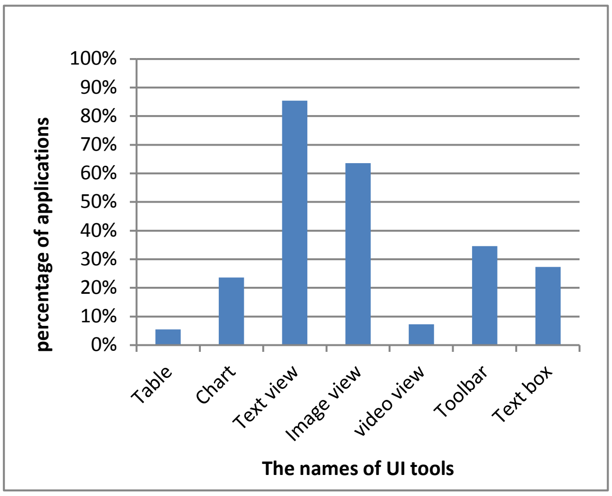


Fig6. Percentage to use tools in applications

4 Applications evaluation

In figure 2 the most user interfaces for mobile applications consists of one page to four pages. This means that the lower the number of pages the more it increased user interaction and affinity and achieve the desired goal of the application such as Samsung Smart Home and Smart pay[28, 48], because enjoy clarity and ease of use and distribution tools are well on the small screen. On the other hand, there are some user interfaces contained 12 pages, this number of pages cause fatigue to the user and make him lose his focus as well as within the context of these applications are Camera Genius 4.2, Yummly and FlatPlayer[24, 30, 49].

The largest proportion of applications that are not using the navigation bar, and this is because of either a user interface for the application of a single page or to refer to the page will be via a button, and this is usually in applications with a small number of pages, such as Racks by the Tracks, Trivia and Simple life[50-52]. In addition, applications that used the navigation bar one, two, three or four such Kuliahmu, Email.me.it, Mail and Free Ringtones [25, 47, 53, 54]. However, the navigation bar in the majority of applications that have a large number of pages. Number of navigation bars depends on the type of application and is designed to think in order to achieve the goal [Fig3].

It seems clear from Figure 4 that most of the applications contained 2 to 21 buttons like Tidean and Half Centric[34, 46]. Meanwhile, some of the applications contained 22 to 63 buttons such Free Ringtones and Camera Genius 4.2[24, 47], the number of buttons is too large and leads to a lack of access to the entire area of the screen. In addition, makes the user lose context. Therefore, the smaller number of buttons in the application provides extra space on the screen and ensuring the user interaction with the application.

From Figure 5 concludes that most applications are not used as menus in applications or a maximum one list such as Messenger, Reportly and Graph[35, 37, 55]. Increase the number of menus in applications confuse the user and does not interact with the application as well lose context like Half Centric, Free Ringtones and FlatPlayer[30, 46, 47]. In addition, the reduction of the menus makes the application simple and easy to use to attract the user.

Although the programming languages provided many of the tools, it was not taken fully exploited. from Figure 6 illustrates that tools used is text view and image view in most applications, that means applications which use these tools, such as MIX, The Porter Beer Bar, Tide and Ideabox[33, 34, 40, 56] are static. However, the authors note the use a few of the table and video view tools. In addition, when used to make a dynamic application this leads to attract the user because the information is always renewed these applications are Smart pay, citizen.tv, Email.me.it[25, 43, 48].

The user interface have to be easy to use from the first interactive user. Functionality to the user must be limited to what the user needs to reach its destination. When they interact with the user interface a matter of expectations should happen.

5 Conclusion

In this article, researchers provide a random sample of user interfaces for mobile applications. Giving an overview of the user interface should be designed for mobile applications, and the problem to be solved to some extent. In addition, the designer should take full advantage of the screen and ensure buttons are placed in a harmonic situation on the screen displays important information only on the surface, and avoid details. Researchers are analysed user interfaces to represent graphically. Then, evaluate the results, which were the most important, the results were as:

• The best applications, which included a one-page to four pages.

• The less use the navigation bar in interfaces provide extra space to exploit due to the small size of the screen.

• The number of buttons whenever less on a single page gives more space to display the information and become more attractive user interface and check a better user interaction.

• Minimizing the menus in applications makes the application easy to use and helps to attract and user interaction, and stays within the context of the application.

• Diversity in the use of tools and good distribution to maintain most important factors to ensure the user interaction with the application.

Finally, designers should be taken of these results in mind when designing interfaces for mobile applications, so as to make the user interface is simple and easy to use, with a nice view to attract the user interacts with it.

6 References

[1] H. B.-L. Duh, G. C. Tan, and V. H.-h. Chen, "Usability evaluation for mobile device: a comparison of laboratory and field tests," in Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, 2006, pp. 181-186.

[2] A. Ginige, M. Romano, M. Sebillo, G. Vitiello, and P. Di Giovanni, "Spatial data and mobile applications: general solutions for interface design," in Proceedings of the International Working Conference on Advanced Visual Interfaces, 2012, pp. 189-196.

[3] C.-H. Leung and Y.-Y. Chan, "Mobile learning: a new paradigm in electronic learning," in Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference on, 2003, pp. 76-80.

[4] Y. Liu, R.-F. Li, C. Xu, and F. Yu, "Design and application of multimedia mobile learning framework," in Machine Learning and Cybernetics, 2007 International Conference on, 2007, pp. 4201-4205.

[5] R. P. Morris and J. J. Tomlinson, "A mobile device user interface with a simple, classic design," Consumer Electronics, IEEE Transactions on, vol. 54, pp. 1252-1258, 2008.

[6] L. Chittaro, "Designing visual user interfaces for mobile applications," in Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems, 2011, pp. 331-332.

[7] B. S. Jin and Y. G. Ji, "Usability risk level evaluation for physical user interface of mobile phone," Computers in Industry, vol. 61, pp. 350-363, 2010.

[8] E. G. Nilsson, "Design patterns for user interface for mobile applications," Advances in Engineering Software, vol. 40, pp. 1318-1328, 2009.

[9] Z. Hussain, M. Lechner, H. Milchrahm, S. Shahzad, W. Slany, M. Umgeher, T. Vlk, and P. Wolkerstorfer, "User interface design for a mobile multimedia application: An iterative approach," in Advances in Computer-Human Interaction, 2008 First International Conference on, 2008, pp. 189-194.

[10] O. W. Bertelsen and C. Nielsen, "Augmented reality as a design tool for mobile interfaces," in Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques, 2000, pp. 185-192.

[11] J. Gong and P. Tarasewich, "Guidelines for handheld mobile device interface design," in Proceedings of DSI 2004 Annual Meeting, 2004, pp. 3751-3756.

[12] W. Park, S. H. Han, S. Kang, Y. S. Park, and J. Chun, "A factor combination approach to developing style guides for mobile phone user interface," International Journal of Industrial Ergonomics, vol. 41, pp. 536-545, 2011.

[13] H. Uzunboylu, N. Cavus, and E. Ercag, "Using mobile learning to increase

environmental awareness," Computers & Education, vol. 52, pp. 381-389, 2009.
[14] D. Park, J.-H. Lee, and S. Kim, "Investigating the affective quality of interactivity by motion feedback in mobile touchscreen user interfaces," International Journal of Human-Computer Studies, vol. 69, pp. 839-853, 2011.
[15] K. Browne and C. Anand, "An empirical evaluation of user interfaces for a mobile video game," Entertainment Computing, vol. 3, pp. 1-10, 2012.
[16] F. Liarokapis and E. Conradi, "User interfaces for mobile navigation," Library Hi Tech, vol. 25, pp. 352-365, 2007.
[17] J. Lumsden and I. Global, Handbook of research on user interface design and evaluation for mobile technology: Information Science Reference, 2008.
[18] J. Huber, J. Steimle, and M. Mühlhäuser, "Toward more efficient user interfaces for mobile video browsing: an in-depth exploration of the design space," in Proceedings of the international conference on Multimedia, 2010, pp. 341-350.
[19] L. F. Motiwalla, "Mobile learning: A framework and evaluation," Computers & Education, vol. 49, pp. 581-596, 2007.
[20] Y.-M. Huang, Y.-T. Lin, and S.-C. Cheng, "Effectiveness of a mobile plant learning system in a science curriculum in Taiwanese elementary education," Computers & Education, vol. 54, pp. 47-58, 2010.
[21] Z. Lu, "A survey of XML applications on science and technology," International Journal of Software Engineering and Knowledge Engineering, vol. 15, pp. 1-33, 2005.
[22] freeware. (2008, 14/02/2014). Gsmart. Available: http://www.freewarepocketpc.net/ppc-download-gsmart-today-plugin.html
[23] G. C. Magda. (2012, 14/02/2014). Mobile App Interface. Available: http://www.behance.net/gallery/Mobile-App-Interface/5432453
[24] PC. (14/02/2014). Camera Genius 4.2. Available: http://www.pcmag.com/slideshow_viewer/0,3253,l=292951&a=292873&po=1,00.asp
[25] email.me.it. (14/02/2014). Video and screenshot tour. Available: http://email.me.it/feature-tour/
[26] PC. (14/02/2014). TrustGo Mobile Security Available: http://www.pcmag.com/slideshow_viewer/0,3253,l=297104&a=297104&po=1,00.asp
[27] R. Inc. (2013, 25/02/2014). Viber iPhone iOS 7 Concept. Available: http://www.behance.net/gallery/Viber-iPhone-iOS-7-Concept/12269245
[28] A. Rahmoun. (2013, 17/02/2014). Samsung Smart Home App Concept. Available: http://www.behance.net/gallery/Samsung-Smart-Home-App-Concept/8532441
[29] Creattica. (22/02/2014). SelfieCam. Available: http://creattica.com/mobile/selfiecam-interface/102225
[30] E. Rahimi. (2013, 16/02/2014). FlatPlayer: The flattest music player you've ever seen! Available: http://www.behance.net/gallery/FlatPlayer-The-flattest-music-player-youve-ever-seen/7730355
[31] e. zuri. (2013, 15/02/2014). Smart Home. Available: http://www.behance.net/gallery/Smart-Home/7412879
[32] A. Szabó. (2013, 15/02/2014). Időkép - weather app redesign. Available: http://www.behance.net/gallery/Idokp-weather-app-redesign/10536853
[33] A. J. Fitzpatrick. (2011, 26/02/2014). Mobile App. Available: http://www.andrewjfitzpatrick.com/?p=1075
[34] M. Spurway. (2013, 17/02/2014). TIDE - SWEET SWELL SURF GUIDE. Available: http://www.behance.net/gallery/Tide/7037109
[35] A. Płatek. (2013, 18/02/2014). Reportly - Time Tracking App. Available: http://dribbble.com/shots/1252296-Reportly-Time-Tracking-App
[36] Creattica. (23/02/2014). Analytics and statistics Available: http://creattica.com/mobile/analytics-and-statistics-app-wip/100844
[37] D. Keller. (2013, 16/02/2014). Messenger app design. Available: http://dribbble.com/shots/1134054-Messenger-app-design
[38] A. Bocharov. (2011, 20/02/2014). Interface design for Quartier Senegalais virtual tour. Available: http://www.behance.net/gallery/Interface-design-for-Quartier-Senegalais-virtual-tour/1421929
[39] T. Media. (2011, 26/02/2014). TaxT. Available: http://www.148apps.com/app/424444782
[40] L. Smith. (2013, 28/02/2014). Mobile Interface Design. Available: http://www.behance.net/gallery/Mobile-Interface-Design/12080535
[41] O. Pineda. (22/02/2014). BurgerQuest. Available: http://creattica.com/mobile/burgerquest/95391
[42] E. Danış. (2013, 15/02/2014). Discovery Channel. Available: http://www.behance.net/gallery/Discovery-Channel/7900623
[43] citizen.tv. (24/02/2014). citizen.tv. Available: http://www.citizentvapp.com/
[44] D. Pires. (2013, 23/02/2014). iOS App Liga Moche. Available:

http://www.behance.net/gallery/iOS-App-Liga-Moche/8385777

[45] Creattica. (22/02/20104). TriplAgent. Available: http://creattica.com/mobile/triplagent-mobile-design/101111

[46] H. Centric. (2012, 22/02/2014). Half Centric. Available: http://www.halfcentric.com/blog/mobile-app-user-inteface-design/

[47] A. Tolstykh. (2013, 18/02/2014). Free Ringtones Available: http://dribbble.com/shots/1169622-Free-Ringtones-Android/attachments/152812

[48] G. C. Magda. (2013, 14/02/2014). Mobile App Interface. Available: http://www.behance.net/gallery/Mobile-App-Interface/5432453

[49] D. Hughmanick. (2013, 15/02/2014). Yummly Mobile. Available: http://www.behance.net/gallery/Yummly-Mobile/10543381

[50] T. d. group. (2013, 14/02/2014). Racks By The Tracks. Available: http://www.timesdigitalgroup.com/gallery-2/#/racks-by-the-tracks-mobile-interface

[51] C. Republic. (20/02/2014). Trivia. Available: http://creattica.com/mobile/iphone-trivia-app-game-ui-design/46709

[52] F. Paul. (20/02/2014). Simple Life Available: http://frinleypaul.com/2012/06/29/simple-life-android-app-user-interface-design

[53] faizalqurni. (2012, 15/02/2014). Kuliahmu App Available: http://faizalqurni.deviantart.com/art/Kuliahmu-App-Mobile-UI-UX-Design-326422013

[54] Michael. (2013, 18/02/2014). Mail for ios 7. Available: http://dribbble.com/shots/1266830-Mail-for-ios-7/attachments/180338

[55] Creattica. (23/02/2014). Graphr. Available: http://creattica.com/mobile/graphr-app/100689

[56] i. ŞENER. (2013, 16/02/2014). ideabox iPhone App. Available: http://www.behance.net/gallery/ideabox-iPhone-App/9860853

An Efficient Soft-Key Interface Provision Method in Cloud Platform

Sanghyun Park[1], Tran Thi Thu Ha[2], Jadhav Yogiraj Shivajirao[3], Jinsul Kim[4]
School of Electronics and Computer Engineering, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 500-757, Republic of Korea
Sanghyun079@gmail.com[1], thuhabkhn@gmail.com[2], yogirajjadhav17@gmail.com[3], jsworld@jnu.ac.kr[4]

Abstract - *In this paper, we propose a software keyboard on a smart mobile environment to provide users with a personalized interface. In order to provide customized user interface efficiently in a cloud environment anytime, anywhere and allowing users to create their own interface directly in a Web platform. Users can use the Web platform in a cloud environment to select the desired interface or they can also create and share their own interfaces. The cloud environment gives users different server platforms, customized interfaces, and a wide variety of applications that can optimize these interfaces. We offer a real time interface to use MongoDB server, complemented with a queuing model algorithm in a cloud environment. Therefore, we provide the interface that efficiently configures the server based on a cloud environment and this is proved through performance analysis and testing.*

Keywords: Cloud Platform, Soft-key Interface, MongoDB, Mobile Interface, Cloud Service

1 Introduction

Recent cloud services are commonly used in our lives. Currently, by using web storage in cloud service IaaS (Infrastructure as a Service) users can store data over a network at any time and can share data from various devices. In addition, it is not only to store the data, but also support viewer functions so user can open file without need software installed in their devices. Smart mobile devices are developing with highly speed that are released in a variety of applications, with a variety of features, according to their personal styles. The smart mobile devices based on 1st version of Android, it is support only for changing the background screen. But, currently, users can customize a variety of features like applications, background, a text message icon on the inside layout, even the default keyboard layout can be changed to suit with their styles. User can download theme from internet and apply into their devices. It is easy to change the background image, but user can't change the keyboard layout. Currently, developers offer few layouts that can be changed by using some application, however, that is limited. Also, interfaces can’t be shared between users. In this paper, we have a wide variety of environments to provide a customized interface efficiently. The interface is based on the user's environment to use with a variety of applications that fit the customized interface. We propose that an interface is provided by the cloud environment, and user create their own layout platform. Further, in order to provide efficient and quick interface MongoDB was used but it has disadvantages. So, this paper proposes a new algorithm is applied into MongoDB server using Queueing Model. Performance analysis and testing is proved through experiments.

2 Related works

The interfaces have two types a hardware interface and a software interface. A lot of users use hardware interface as shown in left right of Figure 1, especially in the QWERTY format used in a PC. On the right side of Figure, the picture shows the keyboard interface which is specially developed for people with disabilities.

Fig 1. Hardware Keyboard

However, currently don’t have variety of hardware keyboard for disability people. If user wants a special hardware keyboard interface, he needs to order separately another one. In recent years, the prevalence of smart mobile devices with touch screens is rising, hence, touchscreen keyboard’s hardware as well as software is able to use. Unlike a hardware keyboard, the software keyboard does not need to create again, by changing only the inner surface of the software, user may be create various software keyboard interfaces. Figure 2 shows the smart mobile devices in a software environment that provides a software keyboard interface. Figure 2 also shows the software keyboard interface with variety of applications. Users can download the theme using their own applications with a smart mobile device. It is easy to change the background, the colors of the interface. However, users cannot create directly the desired interface or variety of interfaces. Users are only offered to make a change in the theme of the interface. So many different people, developers can't provide the desired interface for each person, the existing interface as an interface for the common people, but people with poor vision, the disabled, children and old people who use special type of smart mobile devices, are not provided an easy interface to use. Thus, this paper proposes idea to provide the customized interface based on cloud environment. By which users can custom-tailor the interface with the cloud environment. Users can send and receive variety of interfaces to each other, if there don't have desired interface, users can create their own interface.

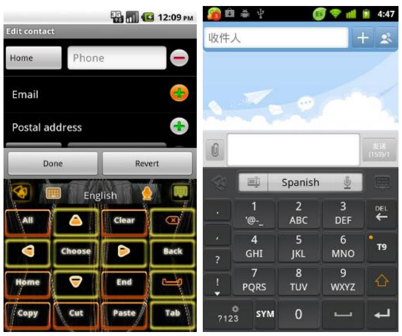


Fig 2. Software Keyboard

3 Cloud system for providing efficient interface

3.1 Cloud-based interface system

We offer cloud-based system structure as shown in Figure 3 order to provide efficiently interface. As shown in Figure 3, the cloud environment inside the network has total four servers with three of them are: "User Data Server", "Interface Server", "Mapping Server", are called "Data Servers" and each server use MongoDB as database system. The user's information is stored in "User Data Server" that provide user with a personalized mobile interface, the user's name, ID, Password, E-Mail, etc. are required to identify the users. "Interface Server" is where to storage interfaces which are created by users and administrators. The administrator can create a wide variety of keyboard interfaced and save them in the "Interface Server" with different interface's ID. Users can access "Interface Server" at any time, to add, modify, or delete a mobile software keyboard interface. Third server is "Mapping Server", consists user' ID and interface's ID to help users join to the cloud system because of MongoDB related to the existing non-relational system, so it is difficult to connect data between "User Server" and "Interface Server". That means, there's no need to access the actual data, just only select information of user's ID and interface's ID to access. Finally, "Interface Provide and Making Tool Server" permits for user and developer connect directly it to send and receive interfaces from cloud environment. The developers can access through a Web browser, and request interface making tool and "Interface Provide and Making Tool Server" provide back interface making tool for developers. When a user want to find new customized keyboard, he send his data request to server and get an new interface which can be shared, deleted, modified by him.

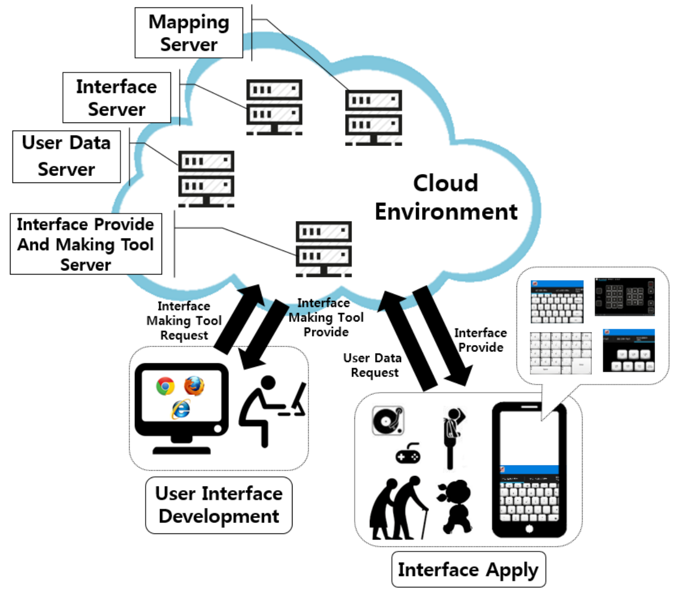


Fig 3. Structure of the whole system

3.2 Platform for customized user interface

In order to create a commonly customized interface for different operating systems, it can be used an interface making tool. The system uses a cloud environment PaaS platform (Platform as a Service) [4]. Users do not need to install a separate SDK and making tools in order to create new interfaces. We have a variety of OS environments, Web-based interface provides a platform for production. Interface making tool based on HTML5 using Web browser to create special interface anytime, anywhere. Figure 4 shows the structure of a Web server to provide the interfaces and making tools. Node.js applications are designed to maximize throughput and efficiency, using non-blocking input/output and asynchronous events. Also, Node.js uses JavaScript for Android-based devices because of the compatibility with each other. Node.js is Google's JavaScript because it uses the constant updates that support faster speeds. Therefore, any data that occurs in a web browser based on the behavior of the Node.js, Node.js via the user data in "Data Servers", the interface data is saved and read. The actual Web-site is offered to the user in the JQuery, HTML5 format that displays interfaces to the user. In addition, JQuery can provide information quickly on the web in a mobile environment. As shown in Figure 4, users can use the smart mobile devices or PCs, interface creation tools can be provided using the JQuery-based platform, and then the user can save a file on the Data Servers through the Node.js [6, 7]. Users can access the interface to request and receive their own interfaces for smart mobile device. A request is made by passing the JQuery data to the Data Servers and with the help of Node.js that access the file. This file is the XML format which is parsed and passed on to the user smart mobiles.

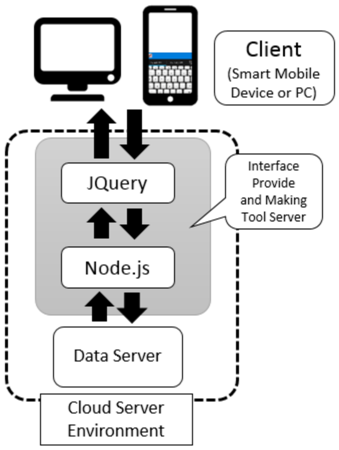


Fig 4. The server structure for providing interfaces and making tools

Fig 5. Interface production platform

Figure 5 shows the interface received platform at the end user through a Web browser. As shown in the illustration, the user can modify the parameters like key size, key location, key height, key width, key color, etc. using the mouse or finger, user can set the value of the key. There are two type of functions. The first function is primarily used in the document, to type a character or function key interface below that fits entered value. The second function is used for video player, music player, games, and fits a wide variety of applications such as an interface for changing video scenes or for increasing volume through the volume up or decreasing volume through volume down. Users can choose different interface as per their favorites or create a new interfaces. Also, users can modify already interfaces which available in server. These interfaces are accessed by interface ID's that can be provided through a cloud environment.

3.3 A method to provide a customized interface

Figure 6 shows how the system provides a customized interface. The users with the smart mobile devices can access the system to receive a customized interface, after user runs the application user has to enter user ID and Password in the Xml format, as shown in Figure 6-1. The requested command is transferred through JQuery, Node.js webservers. The user data is correct or not which is checked by Node.js and search interface information then JSON file is created, saved, the location of the stored Node.js file is transferred to JQuery. The current position values of JQuery, Node.js, are received from the server address, the user ID is sent to the mobile device, the user requested information will be received on the mobile device path which will download the keyboard interface data. The downloaded keyboard interface data can be used in application analysis, users can use the application data to print a keyboard interface.

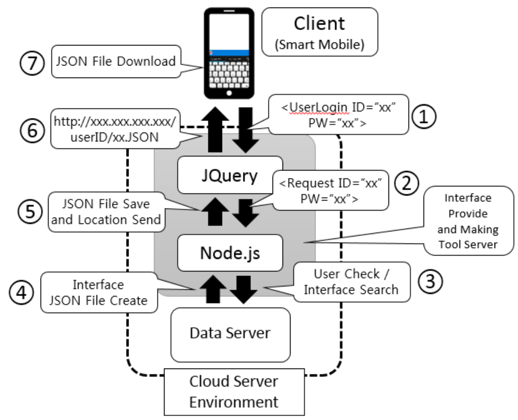


Fig 6. A method for providing a user interface

Internally the data is used to quickly send instructions in Xml format in order to recognize it. XML format has header in the first to identify the commands and understand it in order to read a value that corresponds to an Xml format which sends the request. JSON [8, 9] is one kind of Xml types for the interface transmission of the requested file format. Usually users transfer data in a string with only one interface or single interface, but, when users have multiple data interface and the data is large size hence it can cause lots of damage or data corruption. Therefore, we use JSON file type instead of the string type to send data. The data is downloaded in a single file, hence there is no data corruption. In addition, similar to XML types, the JSON file types are even smaller in size, which makes easier to system to transfer the data easily and quickly. Figure 6-7 shows the user requested interface (see detail in Figure 7)

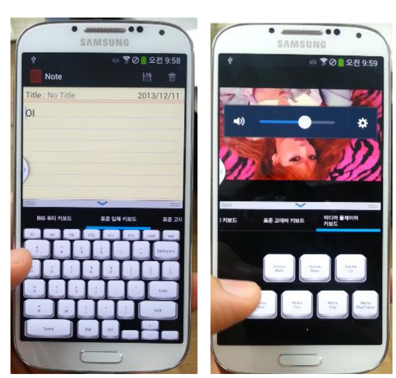


Fig 7. Custom user interface applied to mobile

The system provides a variety of interfaces, As the Sam Sung galaxy S4 supports multi-window feature the application was first tested and run on it as shown in Figure 7, the figure on the left shows a customized interface, that can adjust the volume of the device using the interface provided rather than accessing it through by button as shown in right figure. The top part of the user interface application from the device shows a list of the selected interfaces. For a list of interfaces just slide the screens top portion and user can select the desired interface for a quick use.

3.4 Efficient customized interface for server systems

We tried to explore a variety of methods to provide an efficient interface. Hadoop, which is often used in cloud-based systems that is good to handle big data, but for the system which process data in real time, the result is very slow. In addition, for the distributed system like Hadoop with HDFS is a good preservation method, in this system we are able to add or remove data but the ability to change the data dynamically is not available. Therefore, we used the MongoDB which is capable of performing high-speed data processing in real time. MongoDB can be used to store data as like Hadoop. MongoDB is a data store extension which is highly flexible. It also supports a SQL query function to provide extended functionality and variance. MongoDB is having a variety of functions, such as the indexing spatial information and gathering operation of the MapReduce function in a way that they are built. Recently, there are many smart mobiles on the way so MongoDB provides data languages of JSON type which is in the form of Xml that generates and displays data in the form of a JSON file type. JSON (Java Script Object Notation) is similar to the Xml, but the capacity of the data file is much smaller than the XML which is in characteristic form. MongoDB can be represented as a record in a single complex hierarchical relationship and document-oriented database, in conjunction with object-oriented languages. The system is very suitable from the developer's point of view. In MongoDB there is no need of predefined schema that is different from the SQL-Oracle, hence it is fixed and stores the data flexibly. In order to provide functionality like the index collection, storage, JavaScript, aggregation of fixed size, etc. MongoDB is used. It applies these functions while saving the file, Because of this the server that is different from the SQL-oracle database schema doesn't require changing the schema as it manages automatically. The reading and writing to the database in MongoDB is 3-10 times faster as compared to relational databases. MongoDB is stored in a place called BSON file different from the SQL, the user can save the data via the Web server, MongoDB automatically arranges data to store different types of the servers. It is stored in a place called BSON where user can save data via a Web server, also the user is able to request the data then MongoDB automatically retrieves the data from each of the servers. While saving data, it will be stored in the same manner as Xml format which is provided to the user as a file named JSON as he requests a keyboard interface. The Figure shows the structure of the entire server in the cloud -based environment. When a user requests the data through the device, web server searches the saved data in each of the servers then retrieves user requested data. The BSON is attached to each server as it manages the data stored in each one of them. MongoDB with webserver helps to maintain the data integrity. Thus, it retrieves the data created in the form JSON file that ensures the MongoDB functionality. In this paper, we have used three MongoDB servers that manage data efficiently with high performance.

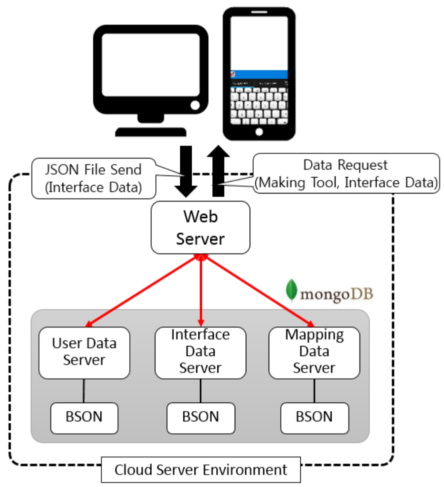


Fig 8. The structure of the server to provide an interface

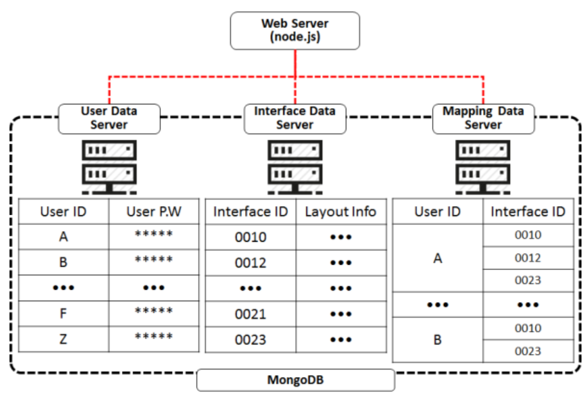


Fig 9. Interface data storage structure

Figure 9 shows the use of Node.js for processing the data of each MongoDB. There is no join function in the MongoDB so we use with Node.js to control MongoDB. Users can use a mobile device or a PC to login, Nods.js check in User Data Server is correct or not. The information of user doesn't fit, Node.js to send error messages. If user information is correct, Mapping Data connect and get interface ID. Mapping Data also prevents the data being duplicated, mapping data can be known depending on whether the user chooses any interface regardless of the entire data interface. Mapping data only has particular ID information as the data does not increase significantly. Thus, node.js uses interface ID and to receive the mapping information in the interface server. Interface Data is made up of values in relation with the interface ID, the size of each key, the position value, the label value, etc. For example, if the user A reads the values 0010, 0012, 0023, Node.js will check for the mapped values and provide the interface related to those values to the user. Finally, the data is created in JSON format which is imported by the user and it is stored in the temporary storage. Node.js sent the user's address to the mobile device and file storage, temporary address of the server. Usually, if users send a string in the stream system with large amounts of data, transfer speed of the user's mobile device may be slow, that may lead to corruption of data, the user by sending the address value, the data can be downloaded to the mobile device, the user's mobile device is able to download a file without data corruption on 3G, 4G, and LTE, in a variety of environments. JSON file is similar to the Xml format, but because of the file size is much smaller than Xml, there should not be an obstacle to transfer the data to the user. Further, it can be seen that it is highly advantages as speed is concerned.

```
WHEN u Request the Keyboard Interface Data THEN
  //mongoose.connect('mongodb://localhost/UserData');
  IF User Check (u) == true THEN
    //mongoose.connect('mongodb://localhost/MappingData');
    i = Mapping Data Search(u)
    WHILE i != NULL THEN
      //mongoose.connect('mongodb://localhost/InterfaceData");
      l = Interface Data Search(i)
    END WHILE
    j = Data File Create (l)
    IF j != NULL THEN
      Data Sending (j)
    END IF
  END IF
END WHEN
```

Fig 10. Algorithm to provide efficient interface

Figure 10 shows the algorithm for providing a user interface, the user is requesting data on mobile devices using the ID and Password will be sent to the server, which is where u refers to the user's ID and Password. User Check function, that determines whether user information is correct or not and finds the user interface by using the Mapping Data function to search users if the ID is detected. *i* is searching interface's ID. By using Interface Data Search function to get information of the interface data. *l* is the interface data, using Data file Create function to create file via JSON. *j* is requested by the user interface, by using Data Sending function to send JSON file for users. In order to provide a customized interface, previously using one server with MongoDB has existing, so now, we use three servers MongoDB make efficient algorithm that determines the performance analysis for Queueing model. Queueing models are used in the data communications and network design theory that speed up the system's behavioral characteristic analysis and design help us to evaluate them. Queueing theory describes how to build a network and related processes for it. The queue first processes the incoming data (First in First out) which is a form of data storage. First, we use M/M/1 Queueing model for performance analysis and M/M/c Queueing model for multiple server analysis. The first M is the probability density for an hour, the second M is the probability density for the time of service, and the last one tells the number of servers, 1 or c. Queueing Model [10, 11] is as follows.

$$\rho = \frac{\lambda}{\mu} < 1 \tag{1}$$

$$L = \sum_{i=1}^{\infty} np(i) = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda} \tag{2}$$

Little's formula ($L = \lambda w$)

$$W = \frac{L}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{1}{\mu-\lambda} \tag{3}$$

$$Lq = \sum_{i=1}^{\infty} (i-1)P(i) = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)} \tag{4}$$

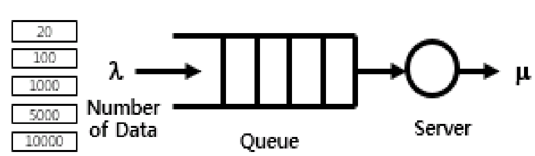
Little's formula ($L = \lambda w$)

$$W_q = \frac{L_q}{\lambda} = \frac{\rho^2}{\lambda(1-\rho)} = \frac{\rho}{\mu-\lambda} \tag{5}$$

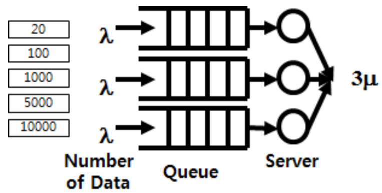
$$W = \frac{L}{\lambda} \tag{6}$$

$$W = \frac{1}{\mu} + \left(\frac{r^0}{c!(c\mu)(1-\rho)^2}\right) p(i) \tag{7}$$

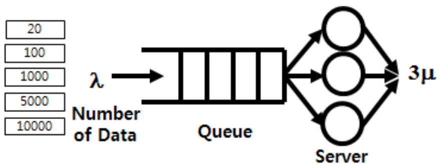
λ is the data, ρ is arrival rate, μ is service rate, Lq is the average queue length, Wq is average waiting time, $P(i)$ (with i: number of users) is probability with conditions. L is the average number in the system is waiting, W is the average time spending in the system (service time), r refers to the average packet in the router. Queueing Model using the standard of equation 5, Equation 6, 7 are performance analysis of M/M/1 and M/M/c Queueing Model. In figure 10, the structure of MongoDB server is designed according to Queueing Model as follow:



(a) Using one server



(b) Algorithm for using three servers and three queues



(c) Algorithm for using three servers and one queues

Fig 11. (a) (b) M/M/1 Queueing Model, (c) M/M/3 Queueing Model

Figure 11, λ refers to the data is inputed, μ refers to the service rates Figure 11-(a) is M/M/1 Queueing Model is applied into one server MongoDB structure. Figure 11-(b) is three Queueing Models M/M/1 and three MongoDB servers, but the algorithm is not applied three servers MongoDB is non-relational in structure to each other because they do not share data with each of the Queue Server will be (M/M/1) + (M/M/1) + (M/M/1). Figure 11-(c) the structure is applied to an algorithm to Queueing Model M/M/ with one queue and three servers. We provide a custom interface effectively using the performance Queueing Model analysis.

4 Experiment results for applied algorithm to customized interface

In this paper, we apply the algorithm to non-relational, MongoDB and after applying the algorithm to the existing relational MongoDB and were tested. First, the reaction rate before the test Queueing Model tests were carried out using a performance analysis. One primary use of the MongoDB server interface for transmitting and efficient algorithms that do not using three MongoDB servers the M/M/1 Queueing Model Algorithm performance analysis was MongoDB because the relational Queue and one server using three M/M/3 Queueing Model to apply a performance analysis. Figure 12 is applied to each of Queueing Model, Performance analysis of the figure shows the result:

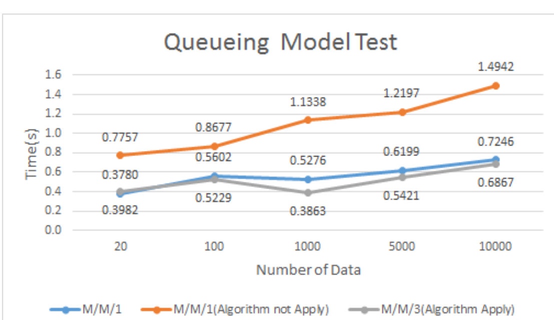


Fig 12. Performance analysis applied to test of Queueing Model

Each data, as shown in Figure 12 to 20, 100, 1000, 5000, 10000 per test were performed. MongoDB M/M/1 servers is used when one of them, M/M/1 algorithm is not applied effectively to transmit data using MongoDB three servers, however, because the non-relational data to have the redundancy. Beets i.e. by 20 per each 20 a total of 60 data is the data to be analyzed. Thus, applying the M/M/1 Queueing Model M/M/1 + M/M/1 + M/M/1, but it can be. Thus, the speed is very slow, as shown in the picture you can see the results were. Finally, the algorithm applied to M/M/3 is a relational data in the 20 because we share to 20/3 data is analyzed and will be divided. When you use one server and apply the algorithm to analyze the performance of three servers, which is not significantly different compared to what you can see. Algorithm with three MongoDB servers is redundant because the data is not shared with each other due to the memory is slowed or more times due to the increased memory, the data has the disadvantage that the reaction rate is slow can be solved. As a result of analysis based on the performance of the actual speed MongoDB tested. Figure 13 shows the actual reaction rates using the MongoDB server as a result of the performance analysis is similar to the test result could be obtained.

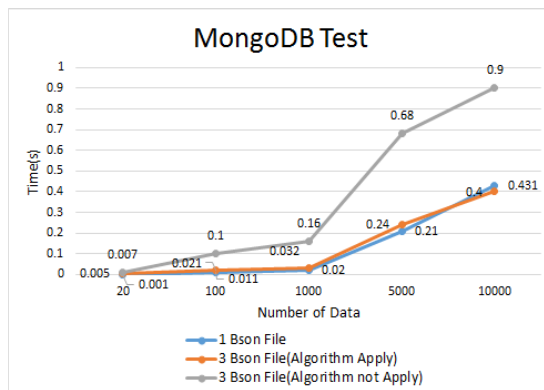


Fig 13. The speed test with different number of MongoDB servers

Data shown in Figure 13 to 20, 100, 1000, 5000, 10000 was tested with data. When one server using MongoDB with MongoDB server to apply the algorithm, the data have been used to analyze the rate that could see almost the same. MongoDB server, but without applying the algorithm to the three, because when you use a non-relational data, have a duplicate. Thus one might have file BSON than about three times more data it can be seen that the reaction rate is coming. When the algorithm was applied to share data between each other, the data is analyzed into three equal parts. The speed is a big difference as shown in the figure, MongoDB server, if you use only one server the redundant data is more than double the amount of data will increase. For example, two people wants to find the keyboard interface A, the keyboard interface provided to each person, selecting and storing a duplicate of the data that had the two. Thus, many people choose to store data in the same interface, the data can be duplicated if more than doubling the number of its conclusion that the two come out. Times or more the data does not need to be duplicated can be a waste of the memory, the data will be processing speed slows down as soon as possible to know. Thus, applying the algorithm to three MongoDB management servers so that the data can be managed effectively and rapidly, which can provide the user interface for performance analysis and test results are demonstrated through experiments.

5 Conclusions

In this paper, we propose a hardware interface and a software interface suitable to the individual providing the interface was in a cloud environment. We can use a cloud platform, users can receive their desired interface, if users do not have the desired interface, and the new interface will be designed. We propose an algorithm for providing efficiently interface that MongoDB server was used, to solve the existing problem of redundancy MongoDB server for managing data using servers in the non-relational data server, the algorithm is applied to make the relational between three servers, through performance analysis and testing can provide a faster data was confirmed. Further, by managing data in a relational interface to a user more effectively been able to provide various interfaces through a cloud environment could be provided anytime, anywhere. In the future research, we will apply our algorithm for different mobile operating systems.

6 Acknowledgements

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1014) supervised by the NIPA(National IT Industry Promotion Agency)

7 References

[1] E. B. Dudin, Yu. G. Smetanin. "A review of cloud computing"; Scientific and Technical Information Processing, Vol.38 No.4., pp.280-284, Oct 2011.

[2] FOTACHE, Marin, Dragos COGEAN. "NoSQL and SQL Databases for Mobile Applications. Case Study: MongoDB versus PostgreSQL"; Informatica Economica, Vol.17 No.2., pp.41-58, April 2013.

[3] Chandrasekaran, Ishwarya. "Mobile Computing with Cloud"; Advances in Parallel Distributed Computing. Springer Berlin Heidelberg, Vol.203., pp.513-522, 2011.

[4] Lorena Gonzalez-Manzano, Gerd Brost, Matthias Aumueller. "Smartphone intelligent applications: a brief review"; Multimedia Systems, pp.1-20, 2013.

[5] Cha, Si-Ho, Yeomun Yun. "Smartphone Application Development using HTML5-based Cross-Platform Framework"; International Journal of Smart Home, Vol.7 No.4., pp.195-201, 2013.

[6] Antunez, Javier Hernandez. "The GPII on Smart Phones: Android"; Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge. Springer International Publishing, pp.58-67, Vol.8514, June 2014.

[7] Heiko Desruelle, John Lyle, Simon Isenberg, Frank Gielen. "On the challenges of building a web-based ubiquitous application platform"; Proceedings of the 2012 ACM Conference on Ubiquitous Computing. ACM, pp.733-736, September 2012.

[8] Chueh, Ting-Fang, and Yong-Yi Fanjiang. "Universal Remote Control on Smartphone"; Computer, Consumer and Control (IS3C), 2012 International Symposium on. IEEE, pp.658-661, June 2012.

[9] Cha, Si-Ho, Yeomun Yun. "Smartphone Application Development using HTML5-based Cross-Platform Framework"; International Journal of Smart Home, Vol.7 No.4., pp.195-201, Jul 2013.

[10] Chen, Yung-Chih, Jim Kurose, and Don Towsley. "A mixed queueing network model of mobility in a campus wireless network"; INFOCOM, 2012 Proceedings IEEE. IEEE, pp.2656-2660, March 2012.

[11] Armony, Mor, and Amy R. Ward. "Fair dynamic routing in large-scale heterogeneous-server systems"; Operations Research, Vol.58 No.3., pp.624-637, Feb 2010.

SESSION

COMMUNICATION SYSTEMS AND TOPOLOGIES + APPLICATIONS AND RELATED ISSUES

Chair(s)

TBA

Efficient Assignment of Packet Cache Region for Traffic Reduction of Multiple Redundant Contents

Yuki Otaka and Akiko Narita
Graduate school of Science and Technology, Hirosaki University, Hirosaki, Aomori, Japan

Abstract - *In recent computer networks, redundant data are often delivered in a short time in large amount, as represented by a live broadcast. We have developed network nodes with packet caches in order to reduce redundant traffic in TCP/IP network. In our system of nodes, An upstream node converts data into a small identifier and a downstream node reconstructs the original data. We have obtained enough reduction rates experimentally for single redundant contents. In this paper, we propose methods for efficient utilization of the packet cache for reducing multiple redundant contents. We implemented the proposed method to a computer network for our experiments and carried out measurements. We improved efficiency of cache utilization and obtained higher reduction rate with a smaller cache.*

Keywords: traffic reduction; network node; packet cache; multiple redundant contents; cache utilization

1 Introduction

In recent computer networks, transmitted data are often redundant. In other words, the same contents are transferred through the same route repeatedly. The redundant data are sometimes generated concurrently as represented by a live broadcast. If we can reduce such redundant traffic, we can utilize limited bandwidth of computer networks efficiently.

Multicast and caching are key methods to reduce redundant traffic. IP multicast is considered the most powerful technique to eliminate redundancy of concurrently generated traffic. Routers at branch points duplicate datagrams with IP multicast. However, restriction in transport layer protocol and equipment discourage the IP multicast from prevailing. In application level multicast, host computers at the end points duplicate data so that it is not constrained by these limitations. Meanwhile, efficiency of bandwidth utilization of it is lower than IP multicast, and overheads are not negligible for constructing a multicast tree in an overlay network. Caching has been implemented to eliminate redundant traffic of data that have been requested repeatedly. One of a typical example is a proxy server. It can inhibit traffic between a real server and it. Caching device for it is usually secondary storage. Unit of caching is a file so that traffic of real time communications is difficult to reduce. Another type of caching is applied in compression and reconstruction by two or more nodes. An upstream node encodes received redundant data and a downstream node decodes it with cached data. It is possible to reduce redundant traffic transparently for the hosts at the endpoints. Furthermore, using primary storage and packet level caching, real time communications are applicable for enough responsibility. That is, the packet level caching is suitable for live broadcast or video conference, which is important use of the recent computer networks.

We have developed the network nodes for reducing redundant traffic by packet caching. We call it TR (traffic reduction) node. Dynamical implementation is feasible for its transparency. OSGi framework is typical to adopt as described in [1]. Moreover, it is also possible to combine other networking function, for example, utilization of lightweight protocol near the low-end host [2]. We obtained enough reduction rates experimentally in simple cases [3]-[5]. We are studying to implement the TR node for more applicative network topologies [6].

It is important for network nodes not to expend resources wastefully. Accommodating cache sharing across multiple redundant contents is significant for effective use of main memory. If redundant traffic consists of two or more contents transmitted concurrently, competitions arise in sharing of cache region across the contents. For efficient use of the cache, such records that will be referred soon should be conserved. Focusing on a single content transmitted by several streams, First in first out (FIFO) is a suitable policy to select a record to replace with new one simply. Relationship between probability of use and order of written is more complicated if multiple contents are transmitted concurrently. The oldest record in the cache may have higher probability of use than a newer record of the other content. In this paper, we present methods for efficient utilization of cache region for multiple redundant contents by grouping redundant streams and restricting cache consumption by a group.

2 Related works

There have been many studies to reduce redundant traffic. [7]-[13] are representative ones. Concepts for traffic reduction in [10] and [13] are very similar to ours. Authors of [10] demonstrated aggressive routing for redundant traffic elimination. They managed redundant data by the method proposed in [7]. A set of Rabin’s fingerprints were identifiers of a cached payload. An algorithm for detecting redundant data in [7] and [10] is more sophisticated than ours. The algorithm is profitable in searching large amount of data. In contrast, for smaller environment temporally and spatially, our easy procedure is considered advantageous. [13] focuses in P2P traffic. Payload of a packet was divided into fixed length blocks and MD5 is used as a hash function tentatively. Only egress routers must store a payload of a compressed packet so that required memory size is expected to be smaller than [10] and our study, instead of less flexibility of applications.

3 Method for traffic reduction

We explain the basic method for traffic reduction in our study based on [5] and modified briefly. Fig. 1 shows a conceptual diagram. The sender generates redundant traffic concurrently. Node U and B reduce traffic cooperatively. They will have the same data in the packet caches shown by Fig. 2 in their main memories. When node U receives a packet, it searches the same data as received one in its cache. If it cannot find the data, it adds the data in the cache. Then it sends the data with request to store in the cache of node D. Otherwise, node U sends identifiers with format in Fig. 3 instead of the received data. Node D reconstructs the original data using cached data and the identifiers. Size of the identifier is much smaller than that of original data so that the nodes can reduce traffic. We implemented our method for the network with Ethernet, IP, and TCP for data link layer, internet layer, transport layer protocol, respectively.

We define a reduction rate as (Dr-Ds)/Dr, where Dr and Ds are amounts of received redundant data and sent data after conversion by the TR node. If all redundant streams are divided at the same offset and transported by the largest frames, node U builds the shortest packets. Then we obtain ideal reduction rate with sufficient cache size as following.

$$\text{ideal reduction rate} = (M-1)\cdot(F-H-S)/(F\cdot M), \qquad (1)$$

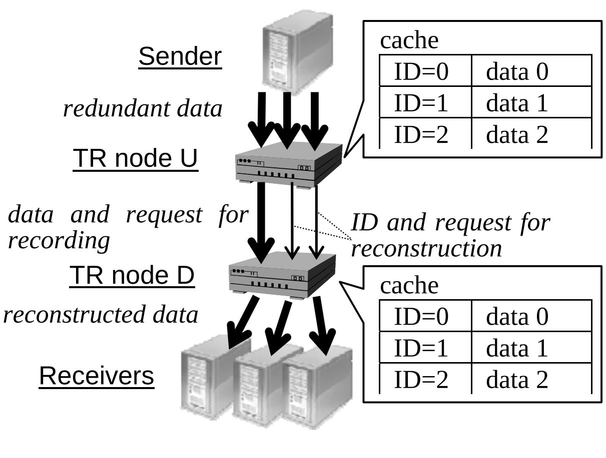


Figure 1. *Redundant Traffic and Basic Method for Reduction.*

where F is maximum frame size, M is number of redundant streams, H is minimum total length of headers of Ethernet, IP, and TCP layer, and S is the minimum length of a data field converted by node U, respectively.

IDf	**property**				**List**
	data	**length**	**IDs**	**flag**	**management**
0					
1					
2					

IDf: Record identifier in the communications cache.
data: Payload of TCP.
length: valid length of the data.
IDs: Stream identifier to which the data belongs.
flag: Valid flag of the record.
List management: Identifiers of forward and backword records of various lists for management.

Figure 2. *Format of the Communications Cache.*

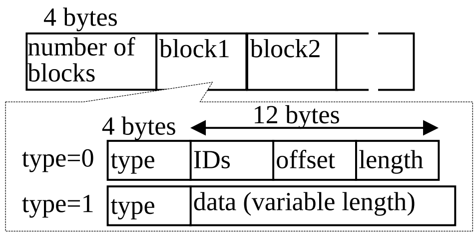


Figure 3. *Format of TCP data converted by the TR node for traffic reduction.*

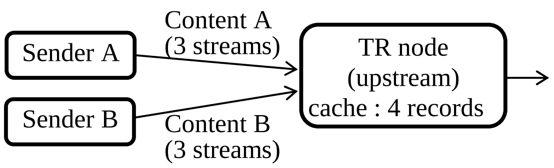
4 Efficient assignment of cache region

4.1 Competition in cache region sharing

If a cache is enough large not to replace a record that will be referred, we can obtain sufficient reduction rate. We call such a record that will be referred “active” in this paper. Required cache size depends on conditions and environments for communications additional to replacement policy. Influential conditions on required cache size are data size that an application program requests to send, transmission rate, window size of TCP decided with performance of hosts and communications bandwidth, deviation of it, and so on.

The optimal condition for saving cache region is that the TR node on upstream always receives packets with completely the same data field (but different destination) successively until it begins to receive the other data. We can see such a condition if a sender transmits fixed size data with low communications rates, and there are no other contents transmitted. In that case, only one record is necessary to accomplish almost ideal reduction rate. In actual case, even if only one content occupies the communications line and application programs always send the same size data, one-record cache is too small. Multiple packet sending before receiving acknowledgement that TCP sliding window allows, resend, and different window increase required cache size. Additionally, the node usually receives not only single redundant content. It is not difficult to distinguish the redundant content from non-redundant one [5]. However, competition in the packet cache is more complicated for multiple redundant contents. The TR node often deletes an active record before it removes inactive ones of the other content. Thus required cache size becomes larger.

The node cannot control these conditions. We should employ appropriate strategy to avoid discarding active record. Consider a simple example of Fig. 4. Sender A and B send content A and B to their 3 receivers, respectively. For simplicity, the senders transmit contents with the same frame size. Capacity of the TR node is 4 records. Segments arrive at the TR node in order of Fig. 4. Strategy for replacing a record influences availability of the packet cache. By optimal (OPT) strategy, the TR node replaces a record that will be referred after the longest term. In an extreme case, the record is inactive. Activity in Fig. 4 means frequency of reference in future. OPT selects a record of the lowest activity. FIFO is suitable for single redundant content except out-of-order arrival at the TR node. In the case of Fig. 4, the TR node sometimes removes active records with FIFO inconveniently in this case.



order	segment (stream)	operation (record number; activity)	
		FIFO	OPT
1	A1 (S1)	Update (1;2)	Update (1;2)
2	A2 (S1)	Update (2;2)	Update (2;2)
3	A3 (S1)	Update (3;2)	Update (3;2)
4	A1 (S2)	Hit (1;1)	Hit (1;1)
5	A2 (S2)	Hit (2;1)	Hit (2;1)
6	A3 (S2)	Hit (3;1)	Hit (3;1)
7	B1 (S4)	Update (4;2)	Update (4;2)
8	B1 (S5)	Hit (4;1)	Hit (4;1)
9	B1 (S6)	Hit (4;0)	Hit (4;0)
10	B2 (S4)	Update (1;2)	Update (4;2)
11	B2 (S5)	Hit (1;1)	Hit (4;1)
12	B2 (S6)	Hit (1;0)	Hit (4;0)
13	A1 (S3)	Update (2;0)	Hit (1;0)
14	A2 (S3)	Update (3;0)	Hit (2;0)
15	A3 (S3)	Update (4;0)	Hit (3;0)
16	A4 (S1)	Update (1;2)	Update (1;2)
17	A5 (S1)	Update (2;2)	Update (2;2)
18	A6 (S1)	Update (3;2)	Update (3;2)

Figure 4. *A Simple example of multple Redundant traffic.*

In practice, it is impossible to know activity of a record exactly. However, we can achieve the optimal replacement in the example focusing attention to characteristics of transmission. For content A, multiple contiguous segments are transmitted sequentially while not for content B. Multiple sending of the frames frequently occurs in transferring bulk data in a short time. Here, we assign spaces for 1 and 3 records for content B and A, respectively, knowing characteristics of the transmission. When the TR node receives a segment of data B2 of stream S4, there are already 3 records of content A. The TR node replaces cache record of numbered 4 conserving the 3 records. It is optimal selection without knowledge of activity. In general, an optimal number is not conclusive for cache records to assign. We propose an algorithm to know preferable size of cache region for redundant contents, that is, number of records to avoid inconvenient replacement.

4.2 Algorithm

For efficient cache utilization, we defined redundant stream group (RSG) and preservation size for it. We prepared the group table and the stream table as shown in Fig. 5 and 6 for management. The RSG is a group of streams that transmit the same content concurrently. The preservation size is number of records that TR node intends not to replace. If Q is a required number of records to obtain sufficient reduction rate for an RSG, number of active records for the RSG is equal to or less than Q. If the newest Q records of the RSG are protected against replacement, active records are conserved. Then Q is possibly an appropriate value for the preservation size. Total sum of the preservation sizes must be equal to or smaller than the packet cache size. Hence, the preservation size may be less than Q. At the same time, there can be more records than the preservation size if cache size is large. Furthermore, we can achieve sufficient reduction rate even if the estimated preservation size is small as shown later.

There are three parts of procedures concerning with record selection for replacement. They are group construction, preservation size calculation, and replaced record selection. The TR node performs these procedures in interrupt operations of packet receiving and timer of interval T. T was set to be one second in the present study. Fig. 7 presents each interrupt operation.

ID*g*	**group size**	**preservation size**
0		
1		
2		

ID*g*: Identifier of the RSG.
group size: number of streams that belongs to the RSG.
preservation size: number of cache records to preserve.

Figure 5. *The Group table.*

IDs	stream	IDg	records	hit count	grouped time
0					
1					
2					

IDs: Identifier of streams transmitted through the TR node.
stream: IP addresses and port numbers of source and destination, and sequence number of record source.
IDg: Identifier of the RSG. Value of -1 means unique stream.
records: Number of cache records that belongs to the stream.
hit count: Frequency of coincidence of data during a timer interruput interval.
grouped time: Life time as a member of an RSG represented by counts of timer intrrupt.

Figure 6. *The Stream table.*

```
--- Operations driven with I/O interrupt ---
Receive a packet carrying data of the stream Sr;
Search the packet cache;
If coincident with data of the stream Sc, {
   Increment hit fields of Sr and Sc in the stream table;
   Set the grouped time at L (L =30 in this paper);
   RSG generation, expansion and integration
      with algorithm 1.1;
} Else select a record to replace with algorithm 3.1
Reconstruct the packet;
Send the packet;
--- Operations driven with timer interrupt ---
For all valid members of the stream table,
   Decrement grouped times of the stream table by 1;
RSG contraction and deletion with algorithm 1.2;
Reserved size adjustment with algorithm 2.2;
Clear hit fields in the stream table;
```

Figure 7. *Operations driven with interrupts.*

4.2.1 Group construction

Identifiers of all streams through the TR node are registered in the stream table. An RSG is composed with registered streams. We call such streams that belong to any RSGs “grouped” and the other streams “unique.” Similarly, records that attribute to the grouped streams “grouped,” and other records “unique.” The TR node generates an RSG, expands it, or integrates two RSGs when the node finds a pairs of redundant streams by the procedure shown in Fig. 8. The TR node contracts an RSG or deletes it when the node finds an improper member as presented in Fig. 9. Deletion of the RSG and its member is necessary to get adequate preservation size. Invalid members in RSGs and actually empty RSGs interfere with calculation of it as described later. Invalid members of the RSG are non-redundant streams or finished ones. The formers are put in the RSG when accidental coincidence occurs in searching cache. Even if only a few bytes in the two streams match accidentally, the TR node regards them as members of the same RSG tentatively. The latters are past streams, whose connections to transmit them have ended. The TR node can distinguish such streams from live and redundant ones using the grouped time in the stream table. When the TR node finds agreement of received data and cached data, it initializes the grouped times of the streams that the data belong to. The initial value of the grouped time is a positive integer L. L was set to be 30 in the present study provisionally. The TR node decrements it in every timer interrupt operation. The grouped time of a truly grouped stream returns to L before it is decreased to 0. In contrast, the grouped time of an ended stream continues to decrease and becomes to be 0. The TR node extracts such a stream from its RSG but does not erase it. Moreover, the TR node deletes the RSG whose group size is 1. If the TR node integrates two RSGs accidentally, affect is resolved when the node extracts the last member of one of the two real RSGs.

```
If both Sr and Sc are unique, {
   Make a new RSG;
   Calculate initial reserved size with algorithm 2.1;
}
Else if only one of Sr and Sc is grouped in an RSG,
   Add the other stream to the RSG;
Else if Sr and Sc grouped in two different RSGs,
   Integrate the two RSGs;

Sr, Sc: See Fig. 7.
```

Figure 8. *Algorithm 1.1: RSG generation, expansion and integration.*

```
For all valid members of the stream table {
   If grouped time value is 0, {
      Break away the stream;
      Decrement size of corresponding RSG;
      If the size becomes 1, delete the group;
   }
}
```

Figure 9. *Algorithm 1.2: RSG contraction and deletion.*

4.2.2 Preservation size calculation

Preservation size calculation is the most significant procedure in the proposed method. There are many approaches for estimation. We employed methods shown by Fig. 10 and Fig. 11 for calculating an initial value and adjusting the preservation size in this paper, because of simplicity and following capability. We use the hit count in the stream table for monitoring deficiency and excess of preservation sizes. The hit count is number of coincidences detected during the timer interrupt intervals. When an RSG does not have enough records to obtain sufficient effect of traffic reduction, the RSG contains a stream with small hit count. On the other hand, the RSG may have too many records if hit counts of all streams of the RSG are large. The

TR node determines whether it should increase or decrease the preservation size of the RSG using the minimum hit count of the streams grouped in the RSG. If the RSG contained an invalid member mentioned in the previous subsection, hit count of the member were always 0 and the preservation size of the RSG continued to increase.

```
A=N/(number of RSGs including new one);
If B=N-(A+(total sum of PR(i)) ≧0, n=A+B;
Else For such i that PR(i)≧A,
   n=PR(i)=(total sum of PR(i))/((number of is)+1);

N: Packet cache size.
PR(i): Preservation size for an RSG labeled i.
n: Preservation size for a new RSG.
```

Figure 10. *Algorithm 2.1: Preservation size calculation for a new RSG.*

```
For each RSG; {
   If hm<hl, classify the RSG into Class 2;
   Else If hl≦hm<hu, classify the RSG into Class 1;
   Else, classify the RSG into Class 0;
}
For Class 0, For each RSG, PR(i)=PR(i)-d;
B=N-(total sum of PR(i));
If B≧d·g2, For each RSG in Class 2, PR(i)=PR(i)+d;
Else If g2≦B<d·g2,
   For each RSG in Class 2, PR(i)=PR(i)+B/g2;
Else For such i that PR(i)≧N/(number of RSGs),
   PR(i)=(total sum of PR(i))/(number of i);

d: Constant integer.
gj: Number of RSGs classified into Class j.
hm: The minimum value of hit counts in an RSG.
hl, hu: Positive integers, hl≦hu,
         hl =800 and lu=1400 in this paper.
N, PR(i): See Fig. 10.
```

Figure 11. *Algorithm 2.2: Preservation size adjustment.*

4.2.3 Replaced record selection

The algorithm is to postpone replacement of an active record. There are two types of an inactive record. One is a record of a non-redundant stream. The other is grouped and too old. The algorithm presented in Fig. 12 selects the inactive record if it exists.

```
For all RSGs,
   If PR(i)<RN(i), the RSG is a candidate group;
Search the oldest record in cache records
   that are unique or member of candidate groups;

PR(i): See Fig. 10.
RN(i): Number of records for an RSG labeled i.
```

Figure 12. *Argorithm 3.1: Replaced record selection.*

5 Performance evaluations and discussion

5.1 Implementation and measurement

We built a computer network for evaluation of the proposed method with machines given in Table I. Basic topology of the network was the same as Fig. 1 but we arranged two sender machines and two receiver machines. Host A and B in the TABLE I sent redundant content A and content B, respectively to the host C and D. Host C and D emulated 5 receivers with processes, respectively. The redundant contents consisted of random numbers. Line speed was 100 Mbps.

We carried out measurements during 3000 seconds with which experimental results were stable. The servers always generated 5 streams per content in the present measurements. Host D was lower-performance computer so that transmission rates of streams reaching host D were smaller. We applied pure FIFO, equal sharing, and the proposed strategy labeled “dynamic” for record replacement in the packet cache. With pure FIFO, which we used in the previous study [2], the TR node grouped all records into only one RSG regardless of contents. With equal sharing, the TR node assigned preservation sizes for RSGs equally for all RSGs instead of the algorithms in Fig. 10 and 11.

TABLE I. HARDWARE AND SOFTWARE SPECIFICATION OF THE EXPERIMENTAL SYSTEM

Function	Specification(CPU/Memory/OS)
TR nodes	AMD Opteron 1210 (1.8 GHz)/1GB /Debian 4.0 (Linux 2.6.18-6-486)
host A /host B	Intel Core i3-3220 (3.30 GHz)/4GB /Ubuntu 11.04 (Linux 2.6.38-16-generic-pae)
host C	Intel Pentium(2.90 GHz)/2GB /CentOS 5.8 (Linux 2.6.18-308.11.1.el5)
host D	Intel Pentium 4 (1.70 GHz) /256MB /CentOS 5.8 (Linux 2.6.18-308.11.1.el5)

5.2 Experimental Results and Discussions

The proposed method directly works out to decrease frequency of replacement. Fig. 13 shows ratio of received packets that caused replacement. With ideal condition, the ratio does not exceed 0.2 for 5 redundant streams. It becomes large if active records are removed. The dynamic strategy decreased replacement rate successfully. Equal sharing gave better rate than pure FIFO for small caches. Performance of the host was lower than that of host C so that content B required smaller region for packets than content A. Equally assigned preservation sizes were appropriate. Improvement with equal sharing means significance of conserving active records discriminating contents even if we cannot calculate the preservation size with high accuracy.

Fig. 14 gives reduction rates. Ideal value of the rate is about 0.75 for 5 redundant streams. Reduction rate reflects decrease of replacement shown in Fig. 13 roughly. Using smaller cache, equal sharing and dynamic give well-improved reduction rate. For small caches, dynamic strategy often assigned equal number of records as preservation size by the applied algorithm. Larger cache size than 600 records, reduction rates with equal sharing are close to those with pure FIFO. From 700-record to 1000-record, half of the cache sizes is considered insufficient to obtain ideal reduction rate for content A. Larger reduction rates are accomplished with dynamic strategy that accommodated preservation sizes between content A and B. For larger cache than 1100 records, reduction rates are almost the same for three strategies. The dynamic strategy is the most advantageous for smaller cache size than 1100. If reduction rate 0.7 is necessary, cache size must be larger than 900 with pure but 700 with dynamic. If reduction rate 0.5 is necessary, cache size must be larger than 500 with pure FIFO but 300 with dynamic.

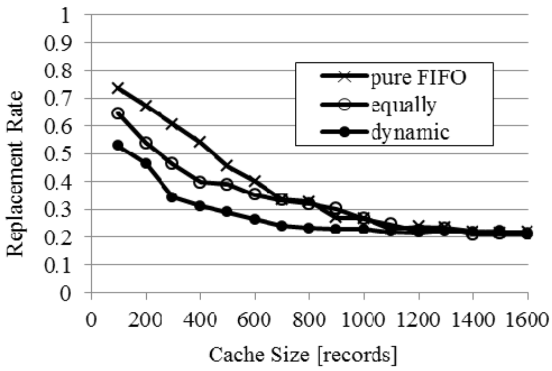


Figure 13. *Replacement Rate.*

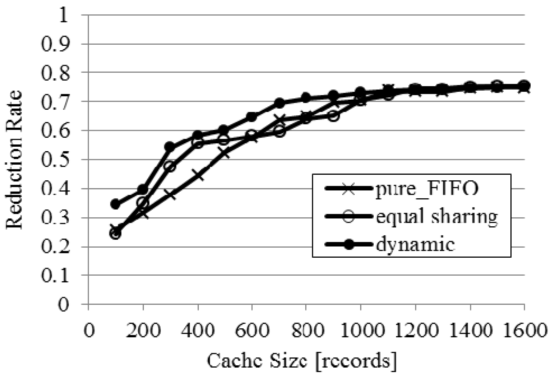


Figure 14. *Reduction Rate.*

Fig. 15 and 16 are number of records in a 700-record cache for 2-redundant contents using pure FIFO and the proposed method. With equal sharing, both contents always consume almost the same number of records. We can realize preservation size effects number of records of each RSG in the packet cache by these graphs. With both of pure FIFO and dynamic, content B uses less cache region than content A on average corresponding to transmission rate. With pure FIFO, recorded sizes for both contents rise and fall with high frequency and wide range. Number of records of content B often becomes larger than that of content A. With dynamic, number of records of content B is always smaller than that of content A except initial state of measurement. The proposed method prevents records of the two contents from chasing off each other successfully.

Fig. 17 shows the preservation size corresponding to Fig. 16. There are much more records of content B than preservation size. On the other hand, preservation size and recorded size of content A are almost the same. Content B used large part of cache region that corresponds to non-preservation size. The preservation size of content A, whose transmission rate was higher, prevented from removing record of content B even if the preservation size of it was very small. Thus, it is not necessary to estimate the preservation size for every RSG accurately but important not to overestimate.

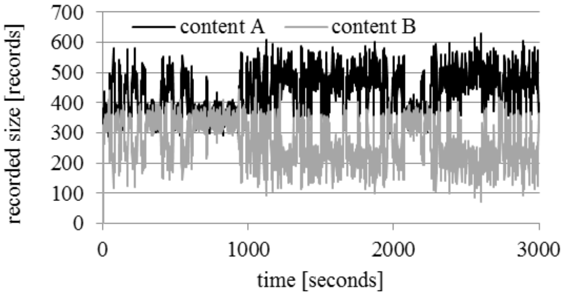


Figure 15. *Number of records in a 700-record cache using pure FIFO.*

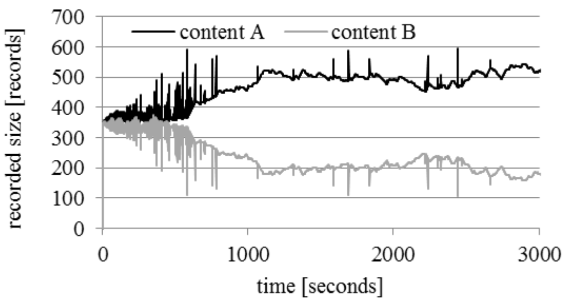


Figure 16. *Number of records in a 700-record cache using the proposed method.*

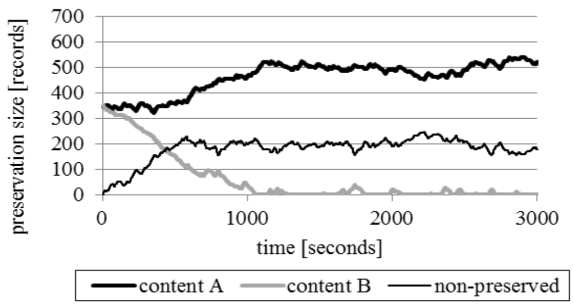


Figure 17. *Preservation size in a 700-record cache using the proposed method*

6 Conclusion and future works

In this paper, we proposed procedures for efficient utilization of cache region for reducing traffic of multiple redundant contents. We defined redundant stream group, estimated required number of cache record to conserve active record for the group, and selected a replaced record omitting the newest ones in the range of the preservation size. We implemented the proposed method to the computer network for experiment and carried out measurements. We decreased frequency of replacement and improved reduction rate using small cache.

In future work, we will investigate subjects to improve efficiency of cache utilization. We consider that the algorithm for preservation size calculation in this paper, especially, threshold to increase or decrease and sharing in small caches, is improvable. We arranged the threshold values used in this paper for the present experiments and the same for all RSGs tentatively. However, they depend on conditions for communications essentially so that we should give different values for each RSG respectively. Otherwise, size, transmission rate and/or mishit rate for RSGs may be valid for the calculation. For small cache, the proposed algorithm abandons not only optimal set of the preservation sizes but also adequate one. The preservation sizes are assigned uniformly except for RSGs with small-converged preservation size. We consider prorating the preservation size depending on transmission rates for RSGs effective. Moreover, for quick convergence of the preservation sizes or applicable following to change of characteristics of communications, we should investigate appropriate amplitude for adjustment of the preservation size. Anyway, better criterion often requires more measures. Complicated procedures generate overhead so that we must be careful to introduce them.

7 References

[1] Jae Woo Lee, Roberto Francescangeli, Jan Janak, Suman Srinivasan, SalmanAbdul Baset, Henning Schulzrinne, Zoran Despotovic, Wolfgang Kellerer. “Net-Serv: Active Networking 2.0”. Proceedings of IEEE International Conference on Communications ICC2011, Jun. 2011.

[2] Koji Funayama and Akiko Narita. “Implementation and Evaluation of a Lightweight Protocol for a Home Network”. Journal of INFORMATION, Vol. 16, No. 7 (B), pp. 5355-5364, Jul. 2013.

[3] Tomoya Shikanai, Kengo Kimura, Sayuri Yamamoto, Yasuyuki Saito, Kenji Ichijo, Akiko Narita, and Yoshio Yoshioka. “Modification and evaluation of a network node for traffic reduction”. Tohoku-Section Joint Convention of Institutes of Electrical and Information Engineers 2010, pp. 138, Aug. 2010.

[4] Yuki Otaka, Yasuyuki Saito, Akiko Narita. “Efficient packet cache utilization of a network node for traffic reduction”. Proceedings of 2012 Annual Conference of Electronics, Information and Systems Society, I.E.E. of Japan, pp. 1867-1868, Sep. 2012.

[5] Yuki Otaka and Akiko Narita. "Efficient Packet Cache Utilization Of A Network Node For Traffic Reduction”. Proceedings of the 2013 International Conference on Internet Computing and Big Data (ICOMP '13), pp. 109-112, Jul. 2013.

[6] Shunsuke Furuta, Yuki Otaka, Akiko Narita. “Design and Evaluation of a Network Node for Traffic Reduction”. The Special Interest Group Technical Reports of IPSJ 2013-CSEC-61, No. 15, pp.-7, May. 2013.

[7] N. Spring and D. Wetherall. “A protocol-independent technique for eliminating redundant network traffic”. Proceedings of the ACM SIGCOMM 2000 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, pp. 87-95, Aug. 2000.

[8] Adam Wierzbicki, Nathaniel Leibowitz, Matei Ripeanu, Rafal Wozniak. “Cache replacement policies revisited: the case of P2P traffic”. Proceedings of the IEEE International Symposium on Cluster Computing and the Grid (CCGrid) 2004, pp. 182-189, Apr. 2004.

[9] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. “I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system”. Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, pp. 1–14, Oct. 2007.

[10] Ashok Anand, Archit Gupta, Aditya Akella. “Packet Caches on Routers: The Implications of Universal Redundant Traffic Elimination”. Proceedings of the ACM SIGCOMM 2008 conference on Data communication, pp. 219-230, Aug. 2008.

[11] Ashok Anand, Vyas Sekar, and Aditya Akella. “SmartRE: an architecture for co-ordinated network-wide redundancy elimination”. SIGCOMM Computer Communication Review, 39 (4), pp. 87-98, Sep. 2009.

[12] Mohamed Hefeeda, Cheng-Hsin Hsu, Kianoosh Mokhtarian. “Design and Evaluation of a Proxy Cache for Peer-to-Peer Traffic”. IEEE Transactions on Computers, Vol. 60, No.7, pp. 964-977, Jul. 2011.

[13] Shu Yamamoto, Akihiro Nakao. “P2P packet cache router for network-wide traffic redundancy elimination”. Proceedings of International Conference on Computing, Networking and Communications (ICNC), 2012, pp. 830-834, Jan. 2012.

SIP over IP VPN: Performance Analysis

A. Asadi Eskandar[1], Mahbubur R Syed[1], and M.B. Zarei
[1] Computer Information Science Department,
Minnesota State University, Mankato, MN, U.S.A

Abstract – *With rapid growth in use of multimedia applications, including IP Telephony (also known as Voice over IP), the demand for security and privacy of communications has significantly increased. Given the fact that IP Telephony utilizes public IP infrastructure, deployment of IP VPN is one approach to protect traffic of interest. However, VPN is presumed to have a negative impact on VOIP performance due to additional packet overheads, authentication, integrity check, encryption, and extra CPU processing involved in the process. In this paper we simulate behavior of a SIP-based VOIP connection running over an IP VPN tunnel. OPNET IT GURU is used as the simulation tool. Few studies examined role of Quality of Service (QoS) parameters such as call signaling and employed VPN protocols, but the role of parameters' arising from network environment, including design issues, were neglected. Bearing this in mind, this research considers several possible scenarios. It compares performance of VOIP calls based on end-to-end delay, delay variation and call setup time for different network configuration (with and without VPN). VPN itself didn't show any negative impact on end-to-end delay. However, combination of network design and VPN show an impact on end-to-end delay.*

Keywords: Session Initiation Protocol (SIP), Voice over IP (VOIP), Virtual Private Network (VPN)

1 Introduction

Voice over IP (VOIP) uses public IP network and its use has been growing rapidly in the past decade. Demand for security and privacy in recent years has prompted various solutions by security researchers and industry role players to assure secure communications over public networks, e.g. Internet. Confidentiality, integrity and authentication of a connection have been considered as a vital security requirement. According to Symantec [1] "registration hijacking" and "eavesdropping" are the two main attacks against SIP-based VOIP communications. Registration hijacking, as its name suggests, involves stealing a VOIP user's subscription and ensuing packets. Eavesdropping, involves intercepting the voice communication streams (RTP packets). In brief, SIP-based VOIP communications are vulnerable to above mentioned attacks because SIP signaling messages are sent in clear text and SIP implementation does not support message integrity check, which makes it easy to modify and replay SIP signaling messages. Thus, in the case of VOIP the session encryption (to avoid issues such as eavesdropping, and man in middle attack), session integrity (to assure VOIP packets are not altered intentionally or unintentionally as they transit over Internet), and protection of related data to a VOIP connection such as packets pertaining to billing system, are the main concerns that needs to be guaranteed.

There exist several well-known and widely used protocols to provide different security and protection features for IP based communications. IPsec (implemented at IP layer), SSL and TLS (implemented at transport layer) are examples of such protocols and mechanisms. One elegant approach towards protecting IP packets is to implement a VPN (Virtual Private Network). Either a site to site VPN or a remote access VPN mode can be designed and implemented to establish an IP tunnel over public IP network. Further details on this are covered in the next section.

Because of the general purpose design of the above mentioned protocols and mechanisms, these may be used for protecting VOIP (and multimedia) packets as well as data traffic. However, VPN may have a negative impact on VOIP performance. It may degrade the quality of VOIP communications due to additional packet overheads, authentication, integrity check, encryption, and extra processing involved in the protection process. Delay (end-to-end delay), jitter (delay variation), packet loss, and call setup time are the main factors that affect the perceived quality of VoIP traffic by the end-user [2, 3]. VPN is presumed to increase the end-to-end delay [3, 4]. The exact amount of delay is dependent on the choice of VOIP signaling protocols (such as H323, and SIP etc.), audio codecs (such as G.711, G.729, G.728, G.726, etc.), related encoding algorithms (such as PCM - Pulse Code Modulation(PCM), Adaptive Differential PCM, etc.), VPN protocols employed (such as PPTP, L2TP, IPsec, etc.), and also parameters arising from network environment [5, 6, 7]. In general, VOIP singling protocols, network Quality of Service (QoS) parameters, and employed security protocols are the broad aspects identified to impact VOIP Quality of Service [5]. In the case of VOIP over VPN, types of services offered to the VPN clients will depend on the choice of VPN protocols (e.g. PPTP, L2TP, GRE, IPSec), having an effect on quality of VOIP traffic. Several researches have addressed VOIP QoS and in particular impact of VPN on VOIP quality performance [3, 5, 8, 9, 10, and 11]. VOIP signaling protocols and choice of audio codec have been examined. However, role of (various types of) VPN protocols and parameters related to the network environment have been neglected or at least have received less attention. Considering this gap, this paper intends to simulate, using OPNET IT GURU network simulation tool, the behavior of a SIP-based VOIP connection, while running over IP VPN tunnel. In particular we set out to compare performance, in terms of end-to-end delay, delay variation and call setup time of a SIP-based VOIP connection running over IP with that of a SIP connection running over IP VPN tunnel. Several possible scenarios, entailing different network designs (in terms of placement of SIP components used in the network) is also created and a comparative analysis is done in order to consider impact of network design on the afore- mentioned issue.

2 Literature Review

Telecommunication standardization bodies such as International Telecommunications Union (ITU-T) have identified a number of factors contributing to QoS for a voice connection. The identified factors include ITU-U voice codecs and algorithm, end-to-end delay, jitter, packet loss, and network design. According to ITU-U guidelines, a voice call facing a delay greater than 150 milliseconds (note: some authors refer to 200ms) and/or a jitter of greater than 20ms is not considered to be of a good quality, and accordingly any voice call facing delay of greater than 300ms and/or a jitter of greater than 50ms is considered to be of a poor quality [12, 13, 14]. Table1 (below) outlines the accepted voice quality measures.

Table 1. Voice Quality Measures [14]

Network parameter	Good	Acceptable	Poor
Delay (ms)	0-150	150-300	> 300
Jitter (ms)	0-20	20-50	> 50

End-to-end delay measures the total amount of time taken for a voice packet to traverse the network path between caller and called entity. Conditions such as poor network capacity (e.g. low bandwidth link, or a low ingress/outbound traffic rate policy associated with a connection), and congestion, will result into voice packet getting delayed, leading into a delay greater than 300ms for voice streams.

Jitter measures the variation from the regular delay time between consecutive voice packets received by the receiver. For example if sender sends consecutive voice packets every certain milliseconds, the receiver is supposed to receive these consecutive packets with roughly similar interval. As mentioned above, 50ms is the maximum tolerable jitter (Note: some authors refer to 75ms).

IP VPN though is an elegant approach towards protecting VOIP packets passing through IP networks, is known to lead to higher delays and jitter in VOIP calls.

Although packet loss, ITU-T codecs & algorithms, and parameters arising from the network environment are factors that play a role in voice QoS, but delay and jitter are considered critical factors that have been examined by most of researches measuring quality of voice calls.

Gouda I. Salama et al. [3] examined “impact of IPsec on the quality of transmitting voice over communication links using OPNET simulator”. Result of their research showed that IPsec results in an increase in packet loss, end to end delay, call setup time, and jitter.

To address QoS for IPsec transmitted packets over IP network, R. Barbieri et al. [9] proposed to down size IPsec encapsulated packets (i.e. the actual IP packet inside IPsec header) by almost 4 bytes using compression. The proposed approach is criticized by Gouda I. Salama et al., [2] for neglecting actual compression time, which in turn will lead to a processing delay that might be even larger than encryption delay.

A.H. Muhamad Amin [5] presented an analysis of VOIP performance measurements using QoS parameters. Congestion was found to negatively impact voice quality parameters such as delay and jitter. VPN showed to have a similar effect on the voice traffic. In a non-ideal network environment, the voice quality parameters showed even worse results compared with an ideal network environment.

Ibrahim S. I. Alsukayti et al. [15] used OPNET modeler to investigate performance of VOIP over VPN running over a BGP/MPLS (Multi-Protocol Label Switching) network. Results suggested that not only a VPN over BGP/MPLS has not a negative impact on VOIP quality but also “positively improves performance of VOIP as compared with its performance over an MPLS network. He also showed that G729A (bit rate = 8 kb/s) is the best choice of voice codec for such scenario (i.e. BGP/MPLS VPN) due to bringing a balance between end-to end delay and bandwidth efficiency. (Note: SIP was the VOIP signal protocol under examination in the study).

2.1 Session Initiation Protocol

Session Initiation Protocol (SIP) is one of the peer to peer VOIP standard protocols defined in RFC 2543 and standardized by the IETF MMUSIC Working Group. This protocol contains initiation, termination and also modification standards for user sessions, which consist of video or audio elements, online games, instant messaging, virtual reality or generally multimedia elements [14]. Table 2 shows SIP massages for a call establishment.

Table 2. SIP Messages [16]

SIP Request Type	Function
INVITE	Session establishment request
ACK	Acknowledgement of receiving INVITE massage
OPTIONS	Capabilities of the server is being queried
BYE	Client asks server to terminate the call
CANCEL	Cancel the request
REGISTER	Registering address with SIP server

Four logical entities defined in SIP are: User Agent (which include: UAC - User Agent Client and UAS - User Agent Server), Registrar, Redirect Server and Proxy Server. User Agent Client (UAC) creates the SIP requests. User Agent Server (UAS) receives a SIP request and responds to it by accepting, rejecting or redirecting the request [17]. Redirect Server's job is to receive requests and reply them with a response message indicating the next place request should be sent (i.e. determining the address of the called device) [17]. For instance, the redirect server would keep track of the user's location and reply a response indicating the location. A SIP Proxy Server acts primarily as a Router by forwarding the SIP request to the next hop. Proxy Severs also play the role of both clients and servers by accepting the users' request and creating a requests on behalf of the users. Proxy Servers are of two types: stateful and stateless. In stateless proxies the reliability of request is not guaranteed and they have the simple function of just forwarding incoming requests to another server. In contrast stateful proxies maintain each transaction's state and include the request and response of that transaction.

The transaction model used in SIP protocol is a request/response model. SIP is considered a text based protocol. A SIP message contains start line, header and body [4]. As mentioned above, the requests are routed to the user's current location using proxy servers. Other responsibilities of proxy servers include user authentication and authorization for services, and call routing policy implementation. Another interesting feature provided by SIP is registration functionality to allow users upload their locations to be later used by proxy servers. SIP runs on top of some transport protocol [3]. The handshake model of the SIP protocol is well illustrated in [2, 19] and also included in figure 1.

TCP and UDP can both be employed for SIP transmission. UDP is more commonly used as the choice of Transport protocol, mainly because of its lower overhead as compared to TCP. Since unreliable UDP transport protocol can be used for SIP massage transmitting, SIP can be responsible for reliability on its own [18].

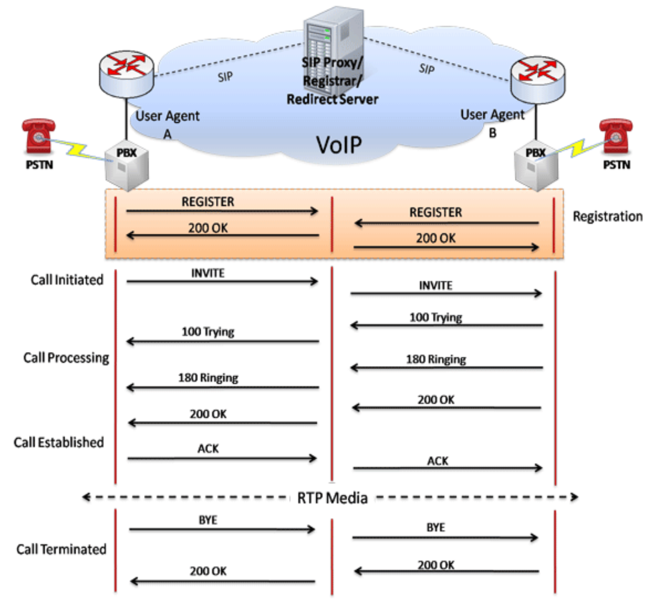


Figure 1. SIP Call Flow [19]

2.2 Virtual Private Network (VPN)

VPNs are responsible for providing a secure connection through a public network. To meet this goal VPN establishes an IP Tunnel- a virtual point to point link between end-nodes, which are separated by a random number of networks in between [9]. VPNs are capable of establishing a secure virtual link between different branch offices. Tunneling makes a virtual lease line (by adding extra headers), and protects private information from being accessed and modified by unauthorized entities throughout the public network [10]. So, tunnel in a VPN is actually a virtual pipe, and responsible for making physical network seems transparent to the packets as they move on their way across the internet. Tunneling can also be referred as the encapsulation process of IP packets into another outer IP header. Figure 2 illustrates a VPN tunnel. VPN uses a dedicated virtual link (i.e. tunnel), for transferring data from source to destination. So, the chance for proxy interruption is minimized. There exist two types of tunnel: permanent and temporary. Permanent tunnels (or static tunnels) are highly resource intensive. This kind of tunnels can be considered too wasteful because they use huge amounts of bandwidth for transmitting not very much data. The disadvantage could be more obvious in a business environment where they are not used 24 hours a day. So in practice a VPN does not utilize a static pipe, instead it uses the more efficient alternative which is dynamic or temporary pipes. These types of tunnel are more flexible considering the fact that they can be established and removed any time needed [12].

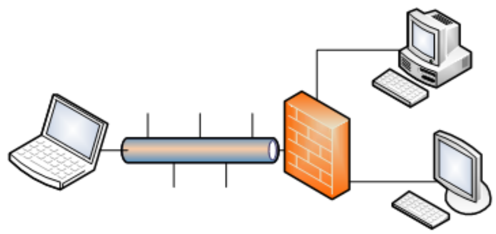


Figure 2. VPN Tunnel

There are two types of VPNs: remote-access and site-to-site. The remote-access VPN establishes a virtual connection between the remote user and the company's internal network. In site-to-site VPN, a virtual connection is established among more than one fixed site via a public network. This type of VPN needs exclusive equipments, because outbound and inbound traffic traverse through peer VPN gateways located in each site. VPN gateways are responsible for encryption/decryption and encapsulation of packets. [11].

Due to complexity of encryption and decryption process in cryptographic algorithms associated with VPN tunnels the performance would become a bottleneck. For this reason, dedicated hardware is proposed as a way to maximize the throughput while the latency gets minimized. In addition, BGP/MPLS VPN technology is being used as a novel VPN approach. In this approach the benefits of MPLS (Multiprotocol Label Switching) technology is combined with security aspects of a VPN. This feature allows service providers to offer VPN service by using their MPLS network for secure delivery of different type of traffic such as voice [8].

3 Methodology

This section describes the simulation tool and processes employed in this research. OPNET IT GURU is used as the simulator in this research. The main research objectives to achieve in this simulation are to:

i. examine the impact of parameters arising from the network and network design on performance of a SIP connection running over IP in terms of end-to-end delay, delay variation and call setup time;

ii. compare the results in i) with a SIP connection running over IP VPN tunnel;

iii. examine the above for several possible scenarios entailing different network designs that include placement of SIP components.

The parameters arising from the network entail a broad range and number of variables and factors, such as network throughput, link capacity and bandwidth, congestion issues, traffic intensity, queuing issue, choice of routing protocols, type of the service provided by the network (e.g. Best effort vs. Guaranteed service), diversity and nature of applications running in the network (Multimedia vs. data), etc.. This research is limited to achieving the simulation objectives stated earlier. Also, to make sure that the primary issue under examination is simulated and experimented with as much control as possible, we kept the possible design scenarios as simple and neat as possible. This would eliminate other possible considerations tied to more complex network design scenarios. Note: EIGRP was employed as the routing protocol in the study.

We created several possible scenarios resembling a simple SIP-based connection over IP backbone with and without VPN. Two IP-phones (using SIP as the signaling protocol) and SIP Proxy servers (one or two - depending on the possible design scenarios) are configured. Proxy server(s) served as the primary source to forward SIP requests and responses. Each scenario was simulated once without VPN and once with VPN in place. For VPN connection, the packets sent from one IP phone to the other one were sent through a VPN tunnel that was established between the routers connected to each IP phone (i.e. ROUTERA and ROUTERB). Figure3 illustrates one of the scenario (named as scenarios 2) used during the experiment.

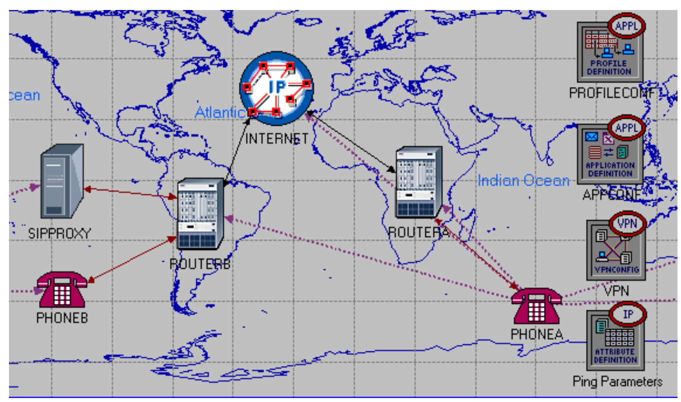


Figure 3. Scenario 2

```
  1. Response Time: 0.31605 seconds

  2. List of traversed IP interfaces:

        IP Address          Hop Delay       Node Name
        ----------          ---------       ---------
         192.0.5.1          0.00000         PHONEA
         192.0.5.2          0.02775         ROUTERA
[label=0 ][exp=0]
         192.0.1.2          0.05293         ROUTERB
         192.0.4.1          0.02116         INTERNET
         192.0.3.2          0.02674         ROUTERC
         192.0.3.1          0.02947         SIPPROXY
         192.0.3.1          0.00001         SIPPROXY
         192.0.4.2          0.02946         ROUTERC
         192.0.1.1          0.02676         INTERNET
         192.0.6.2          0.02114         ROUTERB
[label=0 ][exp=0]
         192.0.5.2          0.05290         ROUTERA
         192.0.5.1          0.02773         PHONEA
```

Figure 4. IP Traffic Flow with VPN connection

Figure 4, illustrates the traffic flow of a ping echo-request and each-response from PHONEA to SIPPROXY. Note: “Compulsory” VPN mode was employed during the experiments, meaning that packets destined to the IP phones would always pass through to the VPN tunnel, no matter if the routing information knew a shorted path to the destination.

Figure 5 shows another scenario (Scenario 3) used in the experiment.

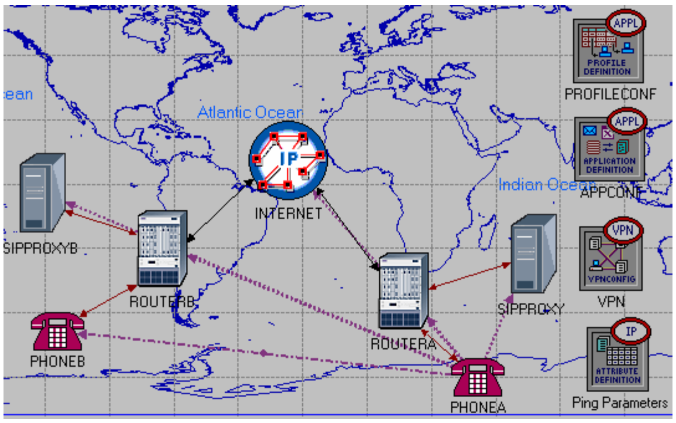


Figure 5. Scenario 3

G.711 is used as the voice codec and SIP as signaling protocol during the experiment. Figure 6 and 7 details the voice codec and voice application definition used in this study.

Name	G.711
Frame Size (secs)	4 msec
Lookahead Size (secs)	0 msec
DSP Processing Ratio	1.0
Coding Rate (bits/sec)	64 Kbps
Speech Activity Detection	Disabled

Figure 6. Voice Codec definition

(Voice) Table

Attribute	Value
Silence Length (seconds)	default
Talk Spurt Length (seconds)	default
Symbolic Destination Name	Voice Destination
Encoder Scheme	G.711
Voice Frames per Packet	1
Type of Service	Best Effort (0)
RSVP Parameters	None
Traffic Mix (%)	All Discrete
Signaling	SIP

Figure 7. Voice Applicaton Definition

In order to generate the VOIP traffic, a VOIP Profile is defined and assigned to PHONEA, which would make constant 2-mintue phone calls to PHONEB during the simulation, which ran for 1 hour. Figure 8 shows the configured profile on PHONEA. PhoneB is configured to support the pre-defined VOIP application as presented in Figure 7, but it is not configured to generate any phone call. This configuration made the controlling and track of VOIP call flows and simulation simpler for later analysis.

Profile Name	VOIP
Applications	(...)
rows	1
row 0	
Name	VOIPAPP
Start Time Offset (seconds)	uniform (5,10)
Duration (seconds)	constant (120)
Repeatability	Unlimited
Operation Mode	Serial (Ordered)
Start Time (seconds)	uniform (0, 0)
Duration (seconds)	End of Simulation
Repeatability	Once at Start Time

Figure 8. VOIP Profile

4 Results

Results of the experiments to achieve the objective, as discussed in the previous section, is summarized in this section.

4.1 Call Setup Time

Findings of this study suggest that VPN will not always lead to higher call setup time, which is different from the findings reported in prior research (e.g. [3]). This research assumed parameters arising from network and network design, will play a role in call setup time of a SIP-based communication. In scenario 2 (fig 3) the IP phones are configured to send the call setup requests to a SIP Proxy server, which is located in the same network segment as PhoneB. In scenario 3 (fig 5) the IP phones are configured to send the call setup requests to their designated SIP Proxy servers, which reside in the same network segment as the IP phones are located. For these two scenarios the VPN show no negative impact on the call setup time. In scenario 1 (fig 9) each of the IP phone is configured to send call setup requests to a SIP Proxy server, which are located in the Internet. In scenario 4 (fig 10) each of the IP phone is configured to send call setup requests to their designated SIP Proxy server, which are located in the Internet. For these two scenarios, the VPN resulted in an increase in the call setup time.

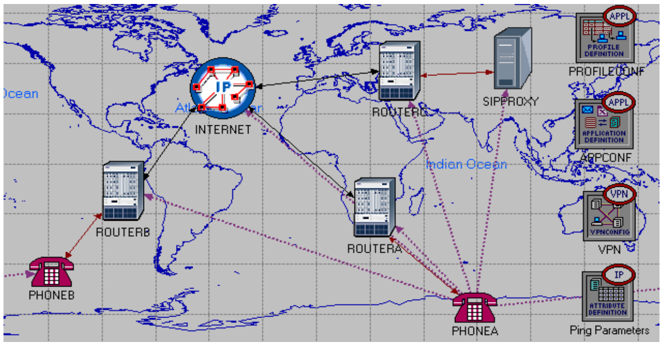


Figure 9. Scenario 1

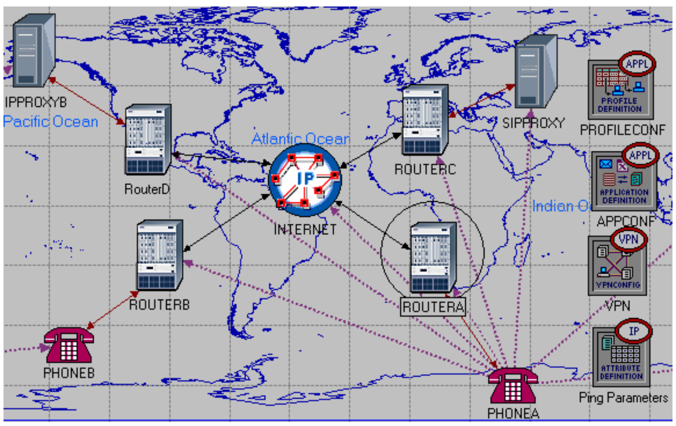


Figure10. Scenario 4

Figure 11 illustrates related call setup time (with and without VPN) for scenario 1 and scenario 4.

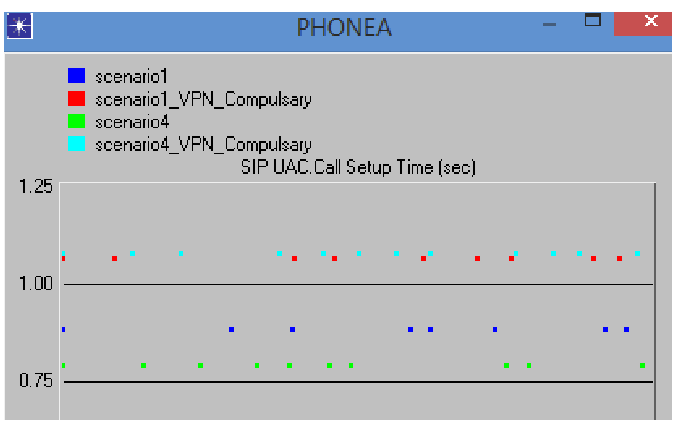


Figure11. Call Setup time for Scenario 1 and 4

Figure 12 presents result of call setup time for all the scnearios. Findings from call setup time show that scenario 3 produced the lowest call setup time amongst all scenarios. Analysis of the call setup time suggests that network design plays an evident role in call setup time and that with a proper design (in this case, proper placement of SIP Proxy server), VPN will not bring a negative impact on the call setup time. Results of call set up time in figure 12 also suggest that using a single SIP Proxy server in the network segment in one of the VPN tunnel end-points is preferred to having the IP phones send their call setup requests to a single designated SIP Proxy server located on the Internet.

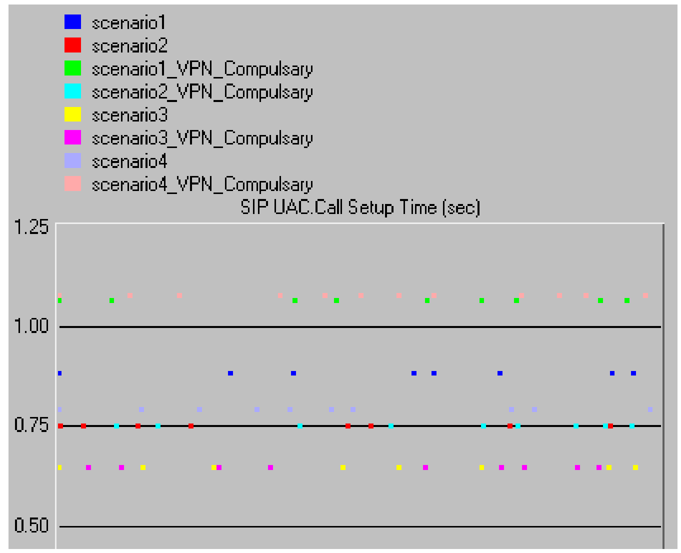


Figure 12. Call Setup Time for All Scenarios

4.2 End-to-End Delay

One significant performance metric for VOIP quality of service is the end-to-end delay issue. Analysis of the end to end delay result, as illustrated in figure 13 and figure 14, suggest that VPN had no negative impact at all on end to end delay. For scenario 1, where each IP phone is configured to send call setup requests to a SIP Proxy server on the Internet, VPN show reduced end-to-end delay compared to when voice streams are forced to go through VPN Tunnel between PhoneA and PhoneB.

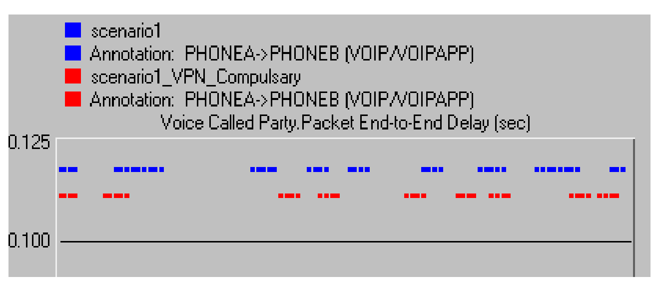


Figure13. End-to-End Delay for Scenario1

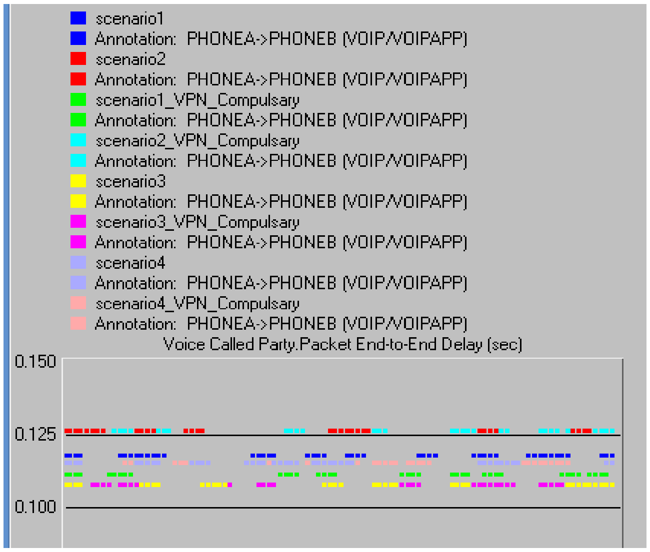


Figure 14. End-to-End Delay

Compartive analysis of end to end delay for all scenarios, as shown in figure 14, suggest that for scenario 3 has the lowest end to end delay. It is interesting to note that VPN do not lead to degradation in VOIP quality in terms of end-to-end delay. Also, VPN reduces the end-to-end delay in conjunction with network design issue. With VPN present in Scenario 2, VOIP end-to-end delay is lower than that of Scenario 4 (with or without VPN). If VPN is not present in Scenario 1, end-to-end delay is higher than that of Scenario 4. A quick reference to Voice Quality Measures as suggested by ITU-T (see table1), suggests that the end-to-end delay results obtained in the experiment are actually good. Therefore, a comparative analysis of end-to-end delay and call setup time indicate that VPN in conjunction with design can actually improve the VOIP performance.

4.3 Delay Varation (Jitter)

This section is allotted to provide a brief review of the Delay Variation results obtained in the simulation experiment. Based on results of figure 15, VPN do not degrade the delay variation.

According to Voice Quality Measures as suggested by ITU-T (see table1), delay variation results obtained in the experiment are good.

5 Conclusions and Future Works

This paper is intended to simulate, using OPNET IT GURU, the behavior of a SIP-based VOIP connection over IP VPN tunnel. Findings of this study suggest that VPN do not always lead to higher call setup time, as opposed to the findings reported in the prior research (e.g. [3]). Analysis of call setup time suggests that network design plays an evident role in call setup time and that with a proper design (in this case, proper placement of SIP Proxy server) VPN will not bring a negative impact on the call setup time. Results of figure 12 (call setup time) also suggest that using a single SIP Proxy server in the network segment in one of the VPN tunnel end-points is preferred than to having the IP phones send their call setup requests to a single designated SIP Proxy server, which is located in the Internet. Compartive analysis of end-to-end delay for all scenarios also suggest that Scenario 3 show the lowest end to end delay.

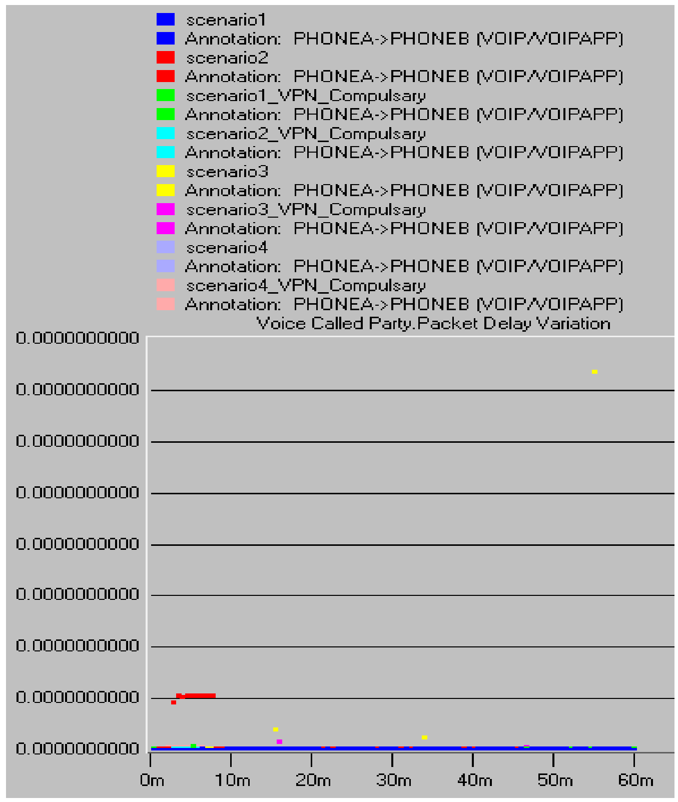


Figure15. Delay Variation

Analysis and findings of this research are different from few related findings e.g. a study by Gouda I. Salama et al., [3] and another study by Muhamad Amin [5] suggested VPN as a factor that increase end-to-end delay. However, findings of this study are aligned with those of Ibrahim S. I et al., [8], where he showed a BGP/MPLS VPN to improve VOIP quality by reducing the end-to-end delay.

Future works may include setting up a network environment scenario in which factors such as network throughput, link capacity and bandwidth, congestion issues, traffic intensity, queuing issue, choice of routing protocols, type of the service provided by the network (e.g. Best effort vs. Guaranteed service), diversity and nature of applications running in the network (Multimedia vs. data) are considered. For instance traffic intensity, which in this study is not set to exceed (or at least closely reach) the link capacities, can be set to a ratio to cover network congestion and throughput issues. A potentially more realistic scenario could also examine end-to-end delay in a network environment in which data packets are sent alongside voice streams.

6 References

[1] Symantec. Two attacks against VoIP. Available: http://www.symantec.com/connect/articles/two-attacks-against-voip

[2] Cisco Systems. VoIP Performance Management and Optimization: Managing VoIP Networks. Available: http://www.ciscopress.com/articles/article.asp?p=1339559&seqNum=7

[3] Gouda I.Salama, M. Elemam Shehab, A. A. Hafez, M. Zaki. Performance Analysis of Transmitting Voice over Communication Links Implementing IPsec in 13th International Conference on AEROSPACE SCIENCES & AVIATION TECHNOLOGY, ASAT- 13, May 2009.

[4] Masqueen Babu. Performance Analysis of IPSec VPN over VoIP Networks Using OPNET. International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 9, September 2012

[5] A.H. Muhamad Amin. VOIP PERFORMANCE MEASUREMENT USING QoS PARAMETERS. The Second International Conference on Innovations in Information Technology (IIT'05).

[6] SANS. Latency and QoS for Voice over IP. Available: https://www.sans.org/reading-room/whitepapers/voip/latency-qos-voice-ip-1349

[7] Cisco Systems. Understanding Delay in Packet Voice Networks. Available: http://www.cisco.com/c/en/us/support/docs/voice/voice-quality/5125-delay-details.html

[8] Ibrahim S. I. Timothy J. Performance Analysis of VoIP over BGP/MPLS VPN Technology in PGNet, 2011.

[9] R. Barbieri, D. Bruschi and E. Rosti, Voice over IPsec: Analysis and Solutions in Computer Security Applications Conference and Proceedings. 18th Annual, Dec2002.

[10] H. Toral-Cruz1, J. Argaez-Xool2, L. Estrada-Vargas2 and D. Torres-Roman. An Introduction to VoIP: End-to-End Elements and QoS Parameters. Available: http://cdn.intechweb.org/pdfs/13379.pdf

[11] Rozalina Dimova, Georgi Georgiev, Zlatko Stanchev. Performance Analysis of QoS Parameters for Voice over IP Applications in a LAN Segment. International Scientific Conference Computer Science'2008

[12] ITU-T. QoS/QoE for VoIP Services. Available: QoS/QoE for VoIP Services

[13] Khaled Alutaibi and Ljiljana Trajković. Performance Analysis of VoIP Codecs over Wi-Fi and WiMAX Networks. Available: http://www2.ensc.sfu.ca/~ljilja/papers/OPNETWORK2012_final_final.pdf

[14] J. Yu and I. Al-Ajarmeh, Call admission control and traffic engineering of VoIP. The Second International Conference on Digital Telecommunications (ICDT 2007) in San Jose, California, USA, July 2007. pp. 11–16

[15] Ibrahim S. I. Timothy J. Performance Analysis of VoIP over BGP/MPLS VPN Technology in PGNet, 2011

[16] Ranganathan, M. Krishna, and L. Kilmartin. "Performance analysis of secure session initiation protocol based VoIP networks" in Computer Communications 26.6 (2003): 552-565.

[17] William Stallings. The Session Initiation Protocol. The Internet Protocol Journal - Volume 6, Number 1.

[18] Shankar R. R. Performance Analysis of Different Codecs in VoIP Using SIP in Mobile and Pervasive Computing, 2008

[19] " VOIP analysis and simulation," http://www.gl.com/mapssip.html. [online].

Modeling of open network reliability including the Internet based on the theory of percolation in two - dimensional and three-dimensional regular and random network structures

D. Zhukov[1], **S. Lesko**[1], **and D. Lobanov**[2]

[1]Department of Technological computer science, Moscow State University of Instrument Engineering and Computer Science, Moscow, Russia

[2]Department of Higher Education Quality, National University of Science and Technology "MISIS", Moscow, Russia

Abstract - *The article describes various models and strategies of blocking network nodes in their address space and physical space. It is shown that the modeling of these processes in the address space only is less informative. The modeling results in the physical space have shown that both in regular two-dimensional (2D) and in three-dimensional (3D) structures the increase of the average number of bonds per one node increases as well the speed of formation of the blocked nodes clusters and the percolation threshold. At that in 3D structures, the speed of clustering and the percolation threshold are higher. The process of the blocked nodes clustering in the 3D structures is faster than in the 2D structures, and the highest efficiency in terms of working capacity have 2D networks with multiple bonds between their nodes. The obtained results allow us to make recommendations for improving the reliability of open networks, including the Internet.*

Keywords: Computer network, blocking nodes, network topology, the percolation threshold

1 Introduction

An important task for ensuring reliable functioning of computer and telecommunication networks as well as transferred information protection is the problem of studying formation of groups of computer networks nodes physically connected by communication channels but blocked (excluded from operation) for some reason. In certain conditions, such groups of blocked nodes can increase in size and form clusters, which can lead to an overall loss of functionality of data network. For example, a cluster can form either when there is a blockage of a backbone node of data network at the regional and city level or belonging to a base station of mobile network as a result of peak load or an overload or when there is a computer virus epidemic in computer networks, which blocks work of different network equipment.

Historically, any information-computer network starting from the city region level has an irregular structure. The brightest example of such a network is the Internet. This is caused by many factors among which we can single out the following: providers having different network and communication equipment, altering number of subscribers with constantly changing connection topology and many others.

When describing topology of network nodes blockage during virus distribution, at present there is an approach, which sees epidemic development as a chain process, reminding Kailey tree with random number of connections by structure [1,2].

A special attention should be paid to a number of papers by R. Pastor – Satorras and A. Vespignani, which consider the task of defining possibility for contamination depending on the distance of the node from the source of infection in the networks with different scale and number of nodes [3-7]. Topological parameters here are the scale and number of nodes but the papers didn't consider variety of network structures.

Common for these papers is the notion of a scale free graph, which can have a different number of nodes. Figure 1a shows an overview of such graph with a total number of nodes of 100.

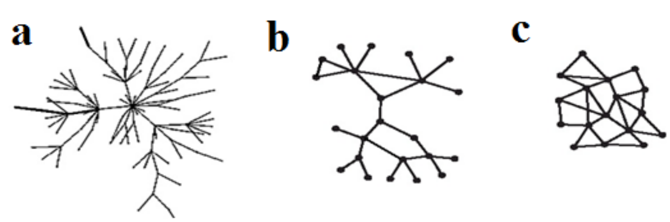


Fig. 1. Overview of scale free graph [2] on Fig. 1a and an overview of infected nodes graph at the later stages of virus epidemic (1b) – the beginning of mutual DDoS attacks, (1c) – the process of mutual DDoS attacks is considerable

The description of virus epidemic topology using scale free graph model produces interesting results. Although, at some stage of the process infected network nodes can start sending copies of viruses to already infected nodes, and the process topology will look as shown on Figures 1b and 1c [1].

Obviously, if the amount of blocked nodes is not too large, there will be an "open" route (way formed by unblocked nodes) between two randomly selected nodes located at a distance. We will refer to amount of blocked nodes at which the network will become nonfunctional as percolation threshold – *l*ower than this value the network will be functional despite the fact that it contains some nodes or their groups (clusters) blocked by viruses. Above the percolation threshold, the whole network turns off and loses its data transfer functionality. There is no "open" way between two randomly selected nodes.

Studying processes of blocked nodes clusters formation and data percolation in networks with different (including random) topology has a lot of scientific and practical importance for development of topology of information-computational networks having high failure-safe feature with the aim of improving their technical and economic as well as operational characteristics and creating new methods and methodologies of computer network and applications protection.

Here two directions can be singled out:

- Research of network kinetics in the *address space* (depending on the number of steps and time) change of the total number of blocked nodes with the application of different algorithms (strategies) of blocking.
- Research of how physical topology of networks influences kinetics of formation of clusters (groups of interconnected closely located nodes) of blocked nodes of different sizes, as well as data percolation at different stages of blockage.

Research in both directions can help develop the most failure resistant architectures and topologies of information-computational networks, as well as methodologies for their defense, including defense of such open networks as the Internet.

Besides, two notions should be distinguished:

- Physical connections between the nodes. Two nodes are considered the neighbors if they have a direct (without intermediate middleman) communication channel.
- Address linkage between the nodes. A virus can send its copy to a randomly selected node with a random IP-address instead of sending it to its physical neighbor.

In the second case, the virus epidemic development topology looks like the Kailey tree (network) with a random number of connections, while in the first the structure of physically connected infected nodes will be more complex and it has almost never been studied.

2 Object and Study Methodology

Choice of network structure for a complex study of topology influence on their reliability in the given study is done based on the fact that assessment of *similarity of real networks* and different theoretical types of topologies on the basis of modeling can help single out a network (or types of networks) with the features closest to those of real networks (for example, the Internet), which is important for analysis of processes happening in the existing networks and ensuring their reliability.

When discussing the issues of topology influencing network reliability, a number of factors can be considered:

- type regular structures symmetry;
- dimension of network structure;
- an average number of connections in one network node.

Based on the above factors, there have been singled out the most typical regular two-dimensional and three-dimensional structures, shown on Fig. 2.

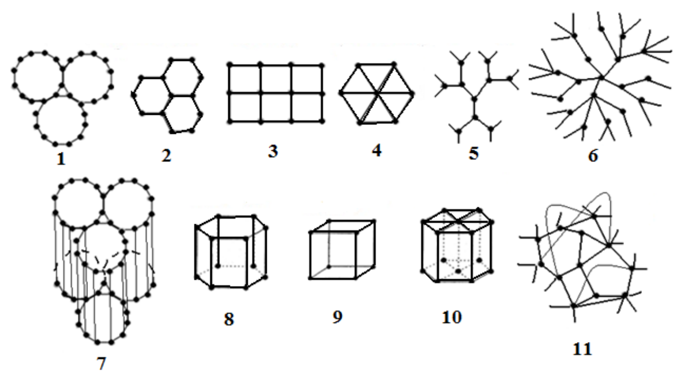


Fig. 2. Graphic representation of studied regular and non-regular two dimensional (2D) and three-dimensional (3D) structures. $3,12^2$ 2D -1, Hexangular 2D - 2, Square 2D - 3, Triangular 2D - 4, Kailey 2D - 5, Random Kailey 2D - 6, $3,12^2$ 3D -7, Hexangular 3D - 8, Cubic 3D - 9, Triangular 3D - 10, Random network with multiple connections - 11.

As random structures, we have studied Kailey networks with a random number of connections and random networks with multiple connections (see Fig. 2). It should be noted that the influence of network dimensions on network reliability should only be considered only for structures with a similar average number of connections per one node, and the influence of an average number of connections should be considered only within one dimension.

As mathematical apparatus in the conducted study, we used the theory of percolation, the basics of which is represented in the works [8-12].

During the modeling, the following assumptions were made: all nodes (10^6) computer network creates a single network with specific topology. Blocking of nodes occurs when infected with a computer virus. The virus can send its copies (10^2) from any node to any other arbitrary node (with probability of infection of $5\cdot10^{-3}$) by selecting its address from the entire set of address space (not necessarily physically connected nearby sites). At the next steps of the epidemic, infected nodes are sending copies of the virus to other nodes in the network, etc. At each step of numerical modeling methods the average cluster size of blocked sites was determined, the infection process was carried out as long as the network did not reach the percolation threshold.

3 Results and discussion

When discussing the processes of clustering, it is important to specify the meaning of topology.

- First, topology can be used to refer to network geometry (triangulated, square, Kailey network, etc. That is quality approach is used).

- Second, in terms of an average number of connections per one node of any network (from regular to random), that is quantity approach.

To discuss the attained results, let us consider how a cluster of 5 blocked nodes (having a structure shown on Fig. 3) is formed in different networks. The dots marked by numbers on Fig. 3 correspond to the networks with the following structures: regular Kailey network (1), 3.12^2 network (2), hexangular (3), square (4) and triangular (5) lattice, 3D structure with the basic 3.12^2 lattice (I), 3D structure with a basic hexangular lattice (II), cubic lattice (III), 3D structure of the basic triangular lattice (IV). Lines on Fig. 3 are to denote the behavior of common dependencies of clustering kinetics for different networks.

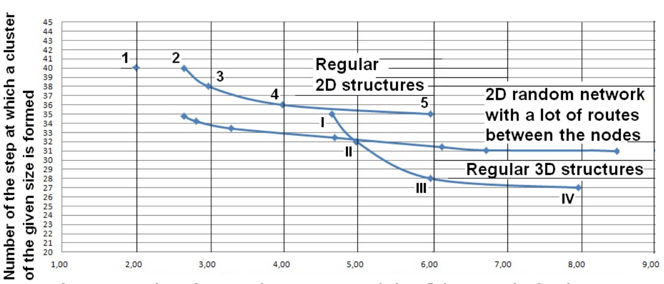


Fig. 3. Dependence of certain size blocked nodes cluster formation from the average number of connections.

It is important to note that studying the processes of small clusters formation is important for networks, which can have blockage of nodes due to their overload as a result of peak loads, long queues of requests and exceeding waiting period. Studying processes of large size clusters formation is more important when nodes are blocked due to viral attacks, when the transition from small to big clusters happens very fast.

As Figure 3 shows, the kinetics of cluster formation in regular 2D structures is slower than in 3D networks (3D сети) and 2D irregular structures. For example, in a regular 2D triangular lattice with the number of connections per one node equal to 6 a cluster of size 5 is formed approximately on the 35th step, and for a random 2D network with the same number of connections – approximately on the 31st. In a square lattice with the number of connections per one node equal to 4, a cluster of size 5 is formed on the 36th step, in an irregular 2D lattice – on the 33rd.

The results on Figure 3 and in Table 1 show that in 3D structures as well as in 2D ones the speed of formation of clusters of different sizes depends on the average number of connections per one node and increases (for clusters of all sizes) with the number of connections increasing.

Table 1.

Network type	Average number of connections per one node in a network of a finite size with a given structure	Number of the step at which the selected cluster size is achieved		
		5 nodes cluster size	500000 nodes cluster size	999900 nodes cluster size
3.12^2 lattice (2D structure)	2.66	40	75	87
Triangular lattice (2D structure)	5.99	35	45	68
Square lattice (2D structure)	3.99	36	47	67
Hexangular lattice (2D structure)	2.99	38	58	73
3D structure with base of 3.12^2 lattice	4.66	35	60	70
3D structure with base of triangular lattice	7.99	27	36	68
Cubic lattice	5.99	28	38	57
3D structure with base of hexangular lattice	4.99	32	50	63

It should be noted that in Table 1 the average number of connections for regular networks, for example, for those with a square lattice structure, equals 3.99, hexangular lattice – 2.99, triangular lattice — 5.99. This is because, although we consider large structures, they are finite; with the boundary playing a certain role (the nodes located at the boundary have a smaller number of connections compared to other nodes). For infinite lattice, the exact theoretical value of an average (per node) number of connections thus equals 4.3 and 6.

Special attention should be given to Keiley network. Regardless of it being regular or having a random number of connections per a single node, the average number of connections per one node of a finite network approaches 2. This is because for any structure of Keiley network the highest number of its nodes is located on the external boundary (leaves of network graph).

For a random network with multiple routes between the nodes the influence of boundary nodes, having a smaller number of connections compared to other nodes, is higher than for networks with regular structures, and the boundary effect for such networks is more substantial.

Modeling of blocked nodes clustering in the networks of different structure (see Table 1 and Fig. 3) leads to the following conclusions:

- For regular 2D structures the speed of formation of clusters of different sizes increases with the increase in the average number of connections per one node (for clusters of all sizes) in the row: regular Keiley network, 3.12^2 network, hexangular, square and triangular lattice. And for regular 3D structures in the raw: 3D structure with a base of 3.12^2 lattice, 3D structure with a base of triangular lattice, cubic lattice, 3D structure with a base of hexangular lattice (hexangular structure).

- With the total average number of connections (see, for example, the 5 to 6 connections area on Fig. 4) per one node being similar, the process of clustering in regular networks with 3D structure is faster than in 2D structures. Random irregular networks with 2D dimension and multiple connections between the nodes take an intermediate position between regular 2D and 3D structures with respect to clustering speed.

- Having a similar spatial dimension of network structure (2D networks), the presence of symmetry makes the process of clustering slower compared to random structures. The increase in dimension from 2D to 3D increases the speed of clustering. In particular, two-dimensional and three-dimensional structures with a similar average number of connections per a node (triangular lattice (2D structure) and cubic lattice (3D structure) – 5.99) can be compared.

- For random 2D networks with multiple routes between the nodes, before at least half the nodes is blocked, the increase in average number of connections per one node increases the process of node blockage (within a homotypic structure).

Figure 4 represent data showing dependence of the amount of infected nodes at which the network of such structure loses its ability to transfer data (percolation threshold) from the average amount of connections per one node in a network of a given structure.

A part of blocked nodes of the whole network at which there is no open data transfer route (unblocked nodes chain) between two randomly selected nodes can be chosen as percolation threshold.

The dotes marked by numbers on Figure 4 correspond to the networks having the following structures: 3.12^2 (1), hexangular (2), square (3) and triangular (4) lattice, 3D structure with a base of 3.12^2 lattice (I), 3D structure with a base of hexangular lattice (II), cubic lattice (III), 3D structure with a base of triangular lattice (IV). The lines on Figure 4 are to define common interdependencies of behavior for the adjusted percolation threshold for different networks calculated for a singular connection.

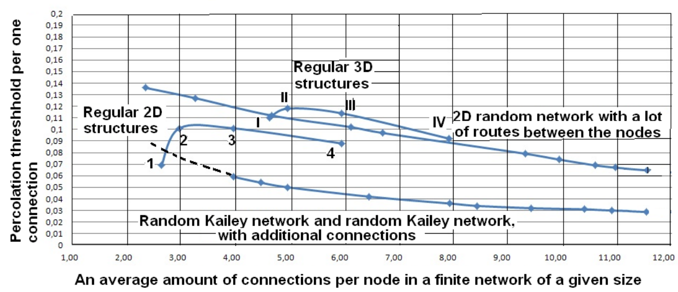


Fig. 4. Dependence of percolation threshold per on node from an average amount of connections per one node in a finite size network (with a given structure)

When constructing a real information-computational network, economic expenses are unlikely to allow to significantly increase the number of connections even between critically important supporting nodes. Obviously, any supporting node can not have more than two connections (otherwise, it will be a hanging node), and the number of connections from 2 to 3 is quite acceptable. Therefore, the most interesting area on Figure 4 is one of a small average number of connections per one node.

Figure 4 shows that with a small average number of connections per one node, the highest percolation threshold per one connection belongs to 2D networks with multiple routes between the nodes (have percolation threshold from 0.5 to 0.9, or per on connection - from 0.12 to 0.06).

An average number of connections per one node in a finite size network with a given structure, as well as the amount of infected nodes at which the network loses its ability to transfer data were defined as a result of numerical experiments.

Random Kailey networks as well as random Kailey networks with additional connections are almost identical in respect to their percolational features.

Increase in dimension leads to increase in an average amount of connections per node. Because of this fact, it is the adjusted values of percolation threshold per one connection that should be compared. The increase in dimension from 2 to 3 with the average number of connections staying the same increases the percolation threshold (reliability) of networks.

Symmetry in the network structure doesn't influence the increase or decrease of percolation threshold, same as change of its dimension (see Fig. 4 area of average number of connections per node from 5 to 8).

These conclusions are very important for potentially increasing the reliability of networks. The Internet is most likely a structure of random Kailey network type with additional connection channels between the nodes. However, due to relatively small changes it can be turned into a random network type structure with multiple routes between the nodes, which can substantially increase its reliability and fault tolerance in data transfer as a result of increasing the value of loss of conductivity (percolation) threshold.

The percolation threshold helps to consider the problem of ensuring reliability from one very important side. To enter a network and block the nodes, the existing computer threats, as a rule, use security vulnerabilities of a certain type (are not multi-vector), which are different for different hardware and software platforms, communication equipment, software, etc., which enable the work of the whole system. If, for example, the Internet is Kailey structure with a random number of connections per one node and additional channels by its topology, then interpolation of percolation threshold behavior per average number of connections of 3 (dotted line on Figure 5) produced the value of adjusted percolation threshold of 0.075, and for the whole structure – 0.225. Therefore, the amount of similar equipment and software for such a network should not exceed 22.5%, and for a random network with multiple connections between the nodes - 39% ($0.13 \cdot 3=0.39$). This can passively prevent loss of network functionality as a whole despite the fact that the functionality and data transfer for some definite nodes and segments can be blocked. Possibly, it is necessary to introduce anti-monopoly ban on part of the used equipment and software market, taking into account percolation threshold for the used networks topology.

4 Conclusions

1. As a result of conducted computational experiments it has been found that in two-dimensional (2D) as well as three-dimensional (3D) structures the increase in an average number of connections per one node increases the speed of formation fo cluster of blocked nodes as well as percolation threshold. In 3D structures the speed of clustering is higher and increase in percolation threshold is faster.

2. Mathematical modeling has shown that with the common similar average amount of connections per one node in regular 3D networks, the process of blocked nodes clustering is faster than in 2D structures. Random networks with 2D dimension take an intermediate position between 2D and 3D structures in terms of clustering speed.

3. Computational experiments have shown that with a small average number of connections per one node, the highest percolation threshold for blocked nodes calculated per one connection, as well as the biggest efficiency in terms of functionality loss, belongs to 2D networks with multiple routes between the nodes.

4. Modeling has shown that increase in dimension from 2 to 3 with the similar average number of connections increases the percolation threshold (reliability and fault resistance) of networks if nodes are blocked.

5. It has been determined that the existence of symmetry in network structure (with the similar average number of connections per node) does not have a substantial influence on increase or decrease of percolation threshold to the same degree as change in dimension.

6. Practical recommendations on how to defend networks from any sort of virus attacks are the following. In case of using equipment and software of similar type their amount should not be higher than 22% for Kailey structure networks with random number of connections and 39% for random network structures with a large number of connections between the nodes.

5 References

[1] Jose Nazario. "Defense and Detection Strategies against Internet Worms"; Artech House, 2004.

[2] Jasmin Leveille. "Epidemic Spreading in Technological Networks". Information Infrastructure Laboratory HP Laboratories Bristol, 2002.

[3] Romualdo Pastor-Satorras, Alessandro Vespignani. "Epidemic dynamics and endemic states in complex networks". Physical Review E, Vol. No 63, 0661171 – 0661178, 2001.

[4] Romualdo Pastor-Satorras, Alessandro Vespignani. "Epidemic spreading in finite scale-free networks". Physical Review E, Vol. No 65, 0351081 – 0351084, 2001.

[5] Romualdo Pastor-Satorras, Alessandro Vespignani. "Epidemic spreading in scale-free networks". Physical Review Letters, Vol. No 86, 3200–3203, 2001.

[6] Romualdo Pastor-Satorras, Alessandro Vespignani. "Epidemics and Immunization in Scale-Free Networks". Wiley-VCH, S. Bornholdt and H. G. Schuster (eds.) Handbook of Graphs and Networks: From the Genome to the Internet, 2002.

[7] Romualdo Pastor-Satorras, Alessandro Vespignani. "Immunization of complex networks". Physical Review E, Vol. No65, 036104, 2002.

[8] Geoffrey Grimmet. "Percolation and disordered systems, in Lectures in Probability Theory and Statistics". Ecole d'Eté de Probabilités de Saint-Flour XXVI-1996, Springer Lecture Notes in Math. no. 1665, ed. P. Bernard, 1997.

[9] Geoffrey Grimmet. "Percolation". Springer-Verlag, 1999.

[10] Michael B. Isichenko. "Percolation, statistical topography, and transport in random media". Rev. Mod. Phys., Vol. No. 64, 961-1043.

[11] Muhammad Sahimi. "Applications of Percolation Theory". Tailor & Francis, 1992.

[12] Vinod K.S. Shante, Scott Kirkpatric. "An Introduction to Percolation Theory". Advances in Physics. Vol. No 85, 325-357, 1971.

Mathematical model of service quality performance's computing of multiservice network

Tulemisova Gulfarida, International academy of business, Almaty, Kazakhstan

Abstract *This article devoted to effectiveness of multiservice networks' functioning in order to provide the quality of these networks. Problem of the probability-time characteristics (PTC) of the network's calculation represents the designers of most interest. Solution of this problem gives you opportunity to: significantly improve network performance, prevent network failures during overload traffic information, to determine the optimal direction of bypass traffic; calculate the optimum quality of service parameters of integrated network. Development of technologies for network management is closely linked to the mathematical modeling of processes and network management elements. For ISDN simulation to determine its threshold or actually parameters of the network, with which it becomes necessary to achieve a particular effect on the control network elements. The article presents the mathematical model calculations PTC, a new method for the formation of nodal load subnet switching channels, a system of equations with respect to the probability of nodal load losses communication channels, obtained a new method for calculating call blocking probabilities between each pair of nodes, we have proved the dependence of these probabilities of blocking probabilities losses for all outgoing directions from the initial node, obtained by the method of calculating the flow channel.*

Keywords: multi-service networks (MSN), mathematical model of probabilistic-time characteristics, methods of hybrid and adaptive switching, adaptive routing, detours traffic, inter-zone connection graph.

1. Introduction

Today, the popularity of the Internet is so vast that the technical capabilities of the World Network do not have time to provide full growing with each passing day the flow of information. Therefore, the task of the qualitative approach to networking is reduced to the task of selecting the topological structure of the network, determining the optimal number of nodes and links of the network to the calculation of its capacity, finding the shortest-path load transfer calculation probability-time characteristics of the network, etc. In the process of analysis and synthesis of networks, the problem of calculating its probability-time characteristics for designers is of greatest interest as a solution to this problem, you can: significantly improve network performance, prevent network failures during overload traffic information, to determine the optimal direction of bypass traffic; calculate the optimal parameters integrated network quality of service.

In multiservice networks delivery of video and voice must be carried out in real time with the need to prioritize in the case of transport network congestion. However, the network industry never focused on real-time network, data delivered in accordance with the capabilities of the network in a specific period of time [1]. An important function of the ISDN is the establishment of bandwidth on demand. Radical extension of bandwidth local area network problems can be solved with QoS, but from an economic point of view in the regional network is not feasible. Therefore, the regional communications network designed to meet the optimization of resource usage for a particular type of traffic. When designing and building a network is widely used technology, which regulates the distribution of costs of various services, provided by the network to transmit information. In such networks in the first place there is the concept of quality of service networks[2].

2. Qualitative problems of multi-service networks.

Quality of Service technology provides the platform necessary for advanced applications, having much more stringent requirements for the width of the provided bandwidth and delay the passage of information in the network. Newer technologies providing quality service depended primarily on various implementations queuing algorithms, which include algorithms "first in - first out" priority queue, fair queue, etc. These algorithms are used by routers and other network devices. But it was impossible to control the continuous flow of traffic. The technology is able to provide the quality of service traffic distribution by category to allow the passage of higher-priority traffic on the network with software settings and regardless of competition from other traffic. Determining when the use of technology is the quality of service is to provide protection for the most priority traffic from various "attacks" by the lower priority traffic, and not just "to use the media on a network".

3. The problem of finding probability-time characteristics of service quality ISDN

This article discusses finding the following statistical parameters:

1) Capacity of branches (bandwidth) subnets and manual CC composed ISDN;

2) The magnitude and nature of the load coming on each hybrid switching node in CC and manual modes and the total load is determined by each group path integral ISDN;

3) The probability of losses in the branches and between each pair of nodes subnet spacecraft, as well as the distribution of these losses on transit routes and nodes in the network, the value of the average delay in the transmission of messages composed subnet CP ISDN.

All of the above statistical parameters during operation ISDN, usually not constant, and changing them is often impossible to predict. So load change and gravity between switching nodes called the commissioning of new units, as well as other factors. In this regard, the modern communications systems a lot of attention paid to the choice of such an algorithm in customer service ISDN, which would take into account the emerging changes in the situation on the network and ensure conservation in a changing set point generalized criterion of quality of service.

4. Method of assessment service quality in subnet CS

This section presents the mathematical model for calculating the parameters of quality of service multiservice network[4]. It uses traffic types multiservice networks using the concept of load transfer detours.

In determining the parameters of quality of service on the subnet QC made the following assumptions:

1) Initial call flows are Poisson;

2) Poisson character is stored as streams for redundant and missing for loads;

3) The system is in a state of statistical equilibrium;

4) System with obvious losses;

5) Does not take into account losses in the switching and control devices;

6) Setup time is zero.

Baseline data to determine the parameters of quality of service are:

1) Structure (location and capacity of the branches);

2) input load for service in BH between the nodes of each pair of nodes;

3) a plan for flows on the network.

4. Probability of losses on branches

In summary, you can get a number of QoS parameters, which were obtained:

1) The magnitude of the total load on each branch;

2) The probability of losses in the branches;

3) The probability of losses between the nodes of each pair of nodes (the ratio of lost load to received);

4) The probability of losses in the middle on the network (the ratio of load lost on the network, received a service);

5) The value of the loads served and lost in each transit node on the network as a whole.

The main parameter here is the probability of losses in the branches, and other parameters can be easily calculated through these values. Calculation of the loss probabilities on the branches in networks with circuitous directions complicated by the fact that the probability of losses in each branch in general depends on the loss probabilities for all other branches.

Score predicted bandwidth and quality of service is a very important stage of designing multiservice networks. Choosing an appropriate traffic model defined selection criteria that use the following parameters:

- Models Incoming calls
- A call blocking,
- Number of sources
- Hold Time.

These developments in the Erlang model is used for trunk groups without repeated calls since callers are redirected or expected low blocking factor. Calculation formula in Erlang traffic models

$$B(c,a) = \frac{\frac{a^c}{c!}}{\sum_{k=0}^{c} \frac{a^k}{k!}} \quad (1)$$

Let’s reform it for more comfortable usage:

$$P(A;v) = \frac{A^v}{v! \sum_{j=0}^{v} \frac{A^j}{j!}} \quad (2)$$

This recurrence relation formula Erlang loss probability calculation and presentation of this formula in integral form. Give proofs of Lemmas, theorems, corollaries. And the formula (1) is transformed into formula (3), (4) and (5).

$$P(A,v) = \frac{A^v}{\sum_{s=0}^{v} \frac{d^s(A^v)}{dA^s}} \quad (3)$$

$$P(A,v+1) = \frac{1}{1 + \frac{v+1}{AP(A,v)}} \quad (4)$$

$$P(A,v) = \frac{A^v e^{-A}}{\int_A^\infty e^{-y} y^v dy} \quad (5)$$

These equivalent formulas much easier to calculate the probabilities of blocking subscriber messages[3].

4.1 Search for missing and excess burden

The magnitude of missing or excess burden depends on the probability of loss $t_i(j)$ of traffic distributed to the branch (ik) .

Distribution of the input load for all subsequent branches of the tree is based on the probability of service, calculated on all the previous branches of the tree paths. In turn, missed branch load $(mn) \in L^u(j)$ is simultaneously input to the load node i.

Finding transit loads at each node of the tree is performed using the following formula:

$$t_i^u(k_1, j) = t_i^u(k_2, j) = \ldots = t_s^u(k_1, j) = r_i^u(j) \qquad (6)$$

Where $r_i^u(j)$ - input load i coming to the site j and designed a node S - the node number of outbound directions. Note that in this case, the input load and stopovers are not fundamentally different from each other. In other words, the transit load input to the process to exclude the cyclic transfer paths. Input burden $r_i^u(j)$ for any node i is distributed over all outgoing directions $(ik) \in L^u(j)$ from it. Above all, since any node in the tree paths, except the first, is included only one branch, then this node, regardless of the choice of outgoing directions there from, the values of transit loads are the same.

In accordance with the criterion of optimality for each branch, a member of one or the other route is determined by some of its weight (cost). The route with the minimum or the maximum weight which is a linear sum of the weights of the branches is optimal according to this criterion, or shortest route. If there are no available machines or free any one branch of the route - the path is considered locked. If all routes set ik blocked, the load is $t_i(j)$ a denial of service. When introduced assumptions subnet circuit switching (CS) is a Markov system with a finite - dimensional phase space E, which state changes occur at discrete points in time corresponding to the time of receipt message to the subnet CS. We assume that the network management U = {Ud, d = 0,1,2 , ...} markov type , then operation of the network can be represented as a controlled markov type process.

4.2 Workarounds direction of load transfer subnet CS

For more efficient use of time channels tract modern automatic switching system located on the nodes, in addition to allow the main ways of communication (the way the first choice) to use workarounds (path next election). For clarity, the above-described use of bypass areas of load transfer CS subnet network is illustrated by the fragment consisting of the *i-th* switching node and out of t_i to neighboring nodes **k1, k2, k3** directions load transfer t_i (Figure 1).

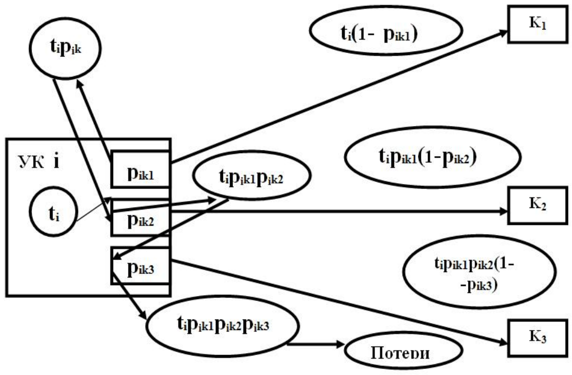


Figure 1 - Using detours transmission nodal load

Ordering of the elements of the set $K_i(j)$ is made in accordance with the choice of node j outgoing direction priority routes in the matrix Mi. Order given by a matrix of load distribution routes:

$$M_i = \begin{array}{c} \\ (ik_1) \\ (ik_2) \\ (ik_3) \\ (ik_4) \end{array} \begin{array}{c} j \\ \begin{bmatrix} \cdots & 3 & \cdots \\ \cdots & 4 & \cdots \\ \cdots & 1 & \cdots \\ \cdots & 2 & \cdots \end{bmatrix} \end{array}$$

In accordance with the elements of this matrix is primarily used branch path of first choice *(ik3*). When it is formed overload excess flow, which is served by a sequence of branches *(ik4), (ik1)* and *(ik2).* Then the set $K_i(j)$ = *{k3, k4, k1, k2}[5].*

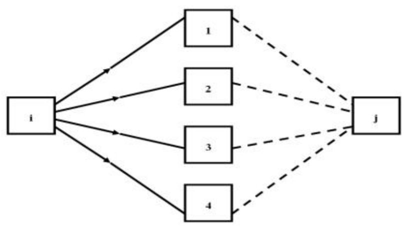


Figure 2 - Count four intermediate network nodes.

Probability of employment service areas (i, $\bar{k}$ *)* is calculated by the formula

$$\varphi_{ik}(j) = \varphi_{i\bar{k}}(j) p_{i\bar{k}}(j) = \prod_{\bar{k} \in \bar{K}_i(j)} p_{i\bar{k}}(j) \qquad (7)$$

Multiplication $\varphi_{ik}(j) * p_{ik}(j)$ is the share of the excess burden on the branch *(ik).*

Probability of losses $\Pi_i(j)$ with load $t_i(j)$ in node j equals to

$$n_i(j) = \prod_{i=1}^{s} p_{ik}(j), \qquad n_i(j) \in (0;1). \qquad (8)$$

When calculating the total load (input at each branch, missed each branch and redundant for this branch) we restrict ourselves to only finding missing branch total load . In turn, through the values $p_{ik}(j)$ can easily determine the total load how each incoming branch and excessive for it.

To select a method for assessing the quality of service on the subnet spacecraft use highly complex adaptive systems, network management, information flows and processes call service, subscribers to share the information. Management system, finding quality deterioration characteristics of various channels, rearranges the order of channels so that these channels will be engaged in the last turn. Likewise was built adaptive control system, which allows reducing any likelihood of loss of calls, the delay time to connect. Quality of service requests for the transfer of information communication system depends on a number of parameters, the main of which usually include: the amount of information to be transmitted, the bandwidth network structural reliability and network survivability, etc.

In fact, as between each pair of nodes i and j are excluded cyclic routes, any node i for a given address is the upstream node relative to k. This means that every path from node i to j, passing through the node k, differs from all the paths between two nodes, k, j branch only *(ik).* In that case losses $P_i(j)$ will be proportionally decreased on amount of branch losses *(ik),* it means $P_i(j) > P_k(j)$. Bandwidth $t_i(j) * P_i(j)$ represents the excess load, and the value **gij** -missed load throughout the paths between nodes i and j. To describe the decentralized algorithms for calculating the main characteristics of data networks using circuitous directions formalized method of optimal allocation of channel resources multiservice network.

4.3 Zonal principle hierarchical addressing and routing

Hierarchical routing algorithm information flow based on the application of the principle of zonal hierarchical addressing and routing. When using a hierarchical addressing zonal address of any node may be represented as a vector

$$A_i = (A_{i1}^1, ..., A_{i,I-1}^1, A_{ii}^1) \qquad (9)$$

Where I - the number of levels of the hierarchical address,

A_{ik}^{j} - the node address in the area of *j-th* level.

When designing hierarchical routing algorithm has adopted the following assumptions:

1) The traffic between the MC levels in a single zone at any level using only the internal path of this zone;

2) The traffic between different zones CC *k- th* level *(k = 1, m - 1)* but belonging to the same zone *(k + 1)-th* level, is sent to the central zone to the CC level, then the path zone *(k + 1)-th* level to the Criminal Code, which is the central area in the *k-th* level, in which the recipient node, and further along the paths zone *k*- level node to the recipient;

3) May allow a direct link between adjacent CC adjacent zones at any level.

Task decomposition network into zones regarded as a goal graph partition

$$G^1 = (X_i^1, Y_i^1) \;\; X_i^1 \subseteq X^1,\, Y_i^1 \subseteq Y^1, \qquad (10)$$
$$i \in I^1 = \{1, 2, \ldots, l_1\},$$

Where, I^1 - number of pieces into which the graph (number of zones l-level).

Partition graph G^1 can be defined by analogy with the partition sets. Collection of pieces of $P\ (G^1)$ is called a partition of the graph $G^1 = (X^1, Y^1)$,

$$\forall\ G_i^1 \in P(G^1)\ [G_i^1 \neq \emptyset],\ i \in I;$$
$$\forall\ G_i^1 \in P(G^1)\ [G_i^1 \neq G_j^1] \Rightarrow X_i^1 \cap X_j^1 =$$
$$= \emptyset \& Y_i^1 \cap Y_j^1 =$$
$$= \emptyset \vee Y_i \cap Y_j = [Y_{ij}], \bigcup_i G_i = G. \qquad (11)$$

In other words, the set pieces of $P\ (G^1) = \{G_i^1, ..., G_{l_1}^1\}$ is a partition of the graph G^1, if any piece of this set is not empty, if for any two pieces of $P\ (G^1)$ the intersection of the set of vertices is empty and the intersection of the set of edges may not be empty , and if the union of all the pieces is exactly equal to the graph G^1.

The expression (11) defines a subset Yij of the set of edges $Y_{ij} \subseteq Y^1$, falling in between the pieces and cut the graph G_i^1 and G_j^1, or in terms of hierarchical addressing , set Yij defines a set of direct links between zones G_i^1 and G_j^1.

Each of the pieces $G_i^1, ..., G_{l_1}^1$, you must select the set of vertices corresponding to

$$X_i^{1,ц} \subseteq X_i^1; X_i^{1,ц} \subseteq S^2 \qquad (12)$$

where the value of S^2 is determined by requirements to network connectivity. (When S^2= 1 there is only one way out of the zone in the Criminal l other areas of the 1-st level, when S^2= 2 - two paths, etc.) Next, we form a graph subnet level 2:

$$G_{1,ц}^2 = (X^2, Y^2), \qquad (13)$$

where $X^2 = \bigcup_i X_i^{1,ц}$, $Y^2 = \subseteq Y_{i,ц}^1$.

Isolation sets $X_{i,ц}^1, \ldots, X_{i,ц}^i$ should be based on the requirements of the connected subnet level 2. Graph G^2 need break apart $G_i^2 = X_i^2, Y_i^2$, $i \in I^2 = \{1,2,\ldots,l_2\}$ and so on up until the next partition to | Ik | = 1. Thus, the resulting partitions m obtain the following relation, specifies the identity of nodes and edges system-oriented graph (SOG) zones of different levels:

$$\{(X_1^1, Y_1^1), \ldots (X_{l_1}^1, Y_{l_1}^1), (X_1^2, Y_1^2), \ldots (X^m, Y^m)\} \qquad (14)$$

Let's call $K_{ij}^s = \sum r_{pq}^s$, $\forall$ (p,q) $\in Y_{ij}^s$ (15)

Thus, K_{ij}^s equal to the total length of all the pieces and the connecting edge G_i^s and G_i^s s of the graph G^s. The connecting edges of all the pieces of the graph on the SIO s-m level

$$K^s = \frac{1}{2} \sum_{i=1}^{l_s} \sum_{j=1}^{l_s} k_{ij} \qquad (16)$$

Total length of all edges of multilevel partitioning

$$K = \sum_{s=1}^{m} K^s \qquad (17)$$

Total length of all edges of multilevel partitioning object of m-level partition graph $G^1 = (X^1, Y^1)$ is to find a set of pieces, the total length of the connecting ribs at all levels meet the specified criteria $K \to \min$.

Let s level graph G^s broken into pieces G_1^s, ..., $G_{l_s}^s$. In accordance with this partition the set of edges Y^s of the graph G^s can be written as

$$Y^s = \bigcup_{i=1}^{l_s} Y_i^s \qquad (18)$$

Then each subset Y_i^s represented as follows:

$$Y_i^s = Y_{i1}^s \cup Y_{i2}^s \cup \ldots \cup Y_{ii}^s \cup Y_{il_s}^s, \qquad (19)$$

Where Y_i^s - a subset of all the edges X_i^s incident to vertices piece G_i^s; Y_{ii}^s - a subset of edges X_i^s connecting vertices subset piece G_i^s together; Y_{ij}^s - a subset of edges connecting pieces G_i and G_j

We call the ratio of the total length of the inner ribs (ribs subsets Y_{ii}^s) to the total length of connecting ribs (ribs subsets Y_{ij}^s) of the partition coefficient $\Delta(G^s)$ graph G^s:

$$\Delta(G^s) = \sum_{i=1}^{l_s} r_{ii}^s / K^s \qquad (20)$$

Multilevel partition coefficient is defined as

$$\Delta(G^s) = \sum_{s=1}^{m} \sum_{i=1}^{l_s} r_{ii}^s / K^s. \qquad (21)$$

This coefficient, as well as the value of K may serve as criterion for evaluating the tiered partition graph. The task refers to the problem of combinatorial and logical type, obtaining an optimal solution which is associated with a lot trying different partition. In order to simplify and reduce the result of the number of options offered to solve the problem using the own algorithm. In this case, each of the $(m - 1)$ steps necessary to solve the problem of partitioning the graph in a traditional setting. However, in this case the number of choices at each level of the partition is large enough. So, for a graph G = (X, Y), | X | = n when splitting into pieces $G_1, \ldots, G_l$ the same dimension n1 = ... = n1 = p the number of options

$$N = \frac{1}{l!} C_n^p C_{n-p}^p \ldots C_p^p = \frac{n!}{l!(p!)^l} \qquad (22)$$

Even for n = 9, l = 3, p = 3, we get N = 1680.

Solve this problem by sorting options are not possible, so it is necessary to use heuristic algorithms.

Let us now determine the size of the zone issues, the choice of metric for determining the length of the edges of the central node and select zones SOG. In this case, since we are considering the problem of separating the control zones SOG on a topological level, we use the results obtained for symmetric networks and symmetric networks with centralization. Selecting an algorithm adaptive routing control within the zone should be based on the coefficient of centralization zone $k_ц = Nc / (N — 1)$, *where* Nc - number of commutation point (CP), related to the central node; N - total number of nodes in the system. The maximum allowable size of the area should be determined by the selected method of adaptive routing. Average packet transfer delay time in the manual mode can be represented as a sum

4.4 Determining the size of the zone

The very choice of adaptive routing control algorithm inside the zone should be based on the coefficient of centralization zone kts = Nc / (N - 1), where Nc - number of the authorized capital, adjacent to the central node; N - total number of nodes in the system. The maximum allowable size of the zone should be determined by the selected method of adaptive routing. Average packet transfer delay time in the manual mode can be represented as a sum of

$$T_a = T_0 + T_c \tag{23}$$

Where T_0 - packet delay time at a fixed routing;

T_c - additional packet delay time, depending on the service traffic.

Given that T_c increases with the size of the zone, it is advisable to ask a certain threshold value and determine the size of the zone based on the relative N_3: $N_3 \longrightarrow \infty$

$$T_c \leq \delta \,. \tag{24}$$

Moreover, since it depends on the choice of control method, the value will depend on the N_3 $\boldsymbol{k_ц}$

$$N_3 = \begin{cases} N_ц & k_ц \leq k_ц^{(2)}; \\ N_д & k_ц > k_ц^{(2)}, \end{cases} \tag{25}$$

Where $k_ц^{(2)}$ - the boundary value of the coefficient of centralization, defined for a given network parameters and δ. Depending on the boundary values of the coefficient of centralization selected centralized or decentralized area.

Also adaptive routing algorithms in digital networks should be implemented algorithms for adaptive control constraint intensity of information flows. The control algorithm is a single set of procedures that control routing and volume (intensity) of the incoming streams of information. Management information exchange can be reduced only to the procedure of flow distribution (routing) in a relatively small traffic, ie, at low intensities in the network flow information.

4.5 Algorithm limits the intensity of flows

In accordance with the decision of the above task and the bandwidth allocation by circuitous paths ISDN load transfer mode commutation point (CP) algorithm restrictions flux intensity can be written as follows:

Step 1. Data Entry: class with multi-channel call (MCC)c, tract integral group(TIG)k;

Step 2. Class definition with(MCC)c,

Step 3. Select (TIG)k according to the routing matrix;

Step 4. If the number of (MCC)c in buffer (TIG)k $l_k^{(c)}$ less than the threshold $l_k^{(c)} < l_k^{(c)}$, then taken to load the information service, i.e. secured her an appropriate buffer. If the condition is $l_k^{(c)} < l_k^{(c)}$, not satisfied, then the load is a denial of service.

This algorithm to determine the flow of information (ADFI) quite simple to implement, but in certain conditions of ISDN, such as a rapid build-up of the load intensity of the upper classes in the commutation point, the restrictions imposed by them, may not be enough.

In order to more quickly and effectively limit the flux intensity in ISDN considers a more complicated algorithm IPR with the exchange of information overload between adjacent CC. The basic idea of the ADFI is that when the number of its own MCC buffer k- *th* TIG i-*th* CC threshold all neighboring j-*th* Commutation point relating to the channel switching (CS) transmitted message about blocking i-*th* CS. After that, the j-*th* CS own MCC who were to be sent to the j- *th* CS are blocked (or forwarded to the workaround). With decreasing values below the threshold all the j-*th* CS messages are sent to withdraw i-*th* block of the commutation point[6]

ADFI consists of two parts. Upon admission to the unit CC, follow these steps:

Step 1. Data Entry: class with MCC, TIG k;

Step 2. Class definition with MCC;

Step 3. Select TIG k in accordance with the matrices routes,

Step 4. If the number of MCC in buffer TIG k does not exceed the threshold of $l_k^{(c)} < L_k^{(c)}$, then go to step 6, otherwise - go to step 9.

Step 5. If $l_k^{(1)} < L_k^{(1)}$ and TIG is not blocked, go to step 6, otherwise - go to step 9.

Step 6. Take a load to the service.

Step 7 Enlarge counter MCC in buffer TIG k per unit : $l_k^{(c)} < L_k^{(c)} + 1$,

Step 8. If TIG k blocked, then finish; if not, send a message about blocking the i-*th* CC and finish.

Step 9. Block CS and finish.

By, the end of service in the i-*th* CC, CS perform the following steps:

Step 10. Decrement counter CS in buffer TIG $l_k^{(c)} = l_k^{(c)}$ *-1*;

Step 11. If $l_k^{(c)} < L_k^{(c)}$ - *d* and k- *th* TIG is blocked, then go to step 3, otherwise finish (parameter d is introduced to ensure the stability of the algorithm);

Step 12. Send to all the j-*th* CC adjacent to the i-*th* of the commutation point, the notice of withdrawal of the lock i-ro CC.

The algorithms ADFI developed effectively control the intensity of flows in ISDN provided a constant ratio between different classes of traffic in CC. $\lambda^{(c_2)}$ = const. I terms of variables $\lambda^{(c_2)}$ ADFI needs, allowing optimal ratio between the quantities $L_k^{(1)}$ - $L_k^{(4)}$ to redistribute - as otherwise the traffic will be blocked by the lower classes if there is sufficient spare capacity buffers CC reserved for the upper classes of traffic. The consequence of this is unjustified decline in performance ISDN.

Let Δt - update interval thresholds; $\tilde{P}$ - Estimate of the average blocking probability MCC i-th interval of the commutation point on $[t_1 - \Delta t, t_1]$; $\tilde{P}^{(c)}$ - Estimate of the average blocking probability MCC class c in the interval in *the* i-*th* CC. The idea of the proposed algorithm is to convert thresholds that, under low load ISDN $\tilde{P} \le P_1$ does not provide any advantages MCC of classes, by increasing a load, ie $\tilde{P} \le P_1$ receives higher priority traffic classes. This algorithm can be included as part of the algorithms described above in the TIG; while periodically at intervals Δt , perform the following steps .

Step 1. Detect mode ISDN, for which to calculate $\tilde{P}$ the interval $[t_1 - \Delta t, t_1]$ - the current time.

Step 2. Secure another traffic class Tc and perform pp 3-5.

Step 3. If | $\tilde{P}^{(c)}$ - $\tilde{P}$ | < ε, where ε a certain threshold value , then go to step 4 , otherwise - to step 5 .

Step 4. If $\tilde{P}^{(c)}$ - $\tilde{P}$ $\ge$ ε - then increase the threshold $L^{(c)}$: $L^{(c)} = L^{(c)} + l_0$.

Step 5 .If $\tilde{P}$ - $\tilde{P}^{(c)} \ge$ then reduce the threshold: $L^{(c)}$: $L^{(c)} = L^{(c)} - l_0$ (l_0 - increment threshold).

Step 6. If considered all traffic classes, then finish, otherwise - go to p.2.

IPR developed control algorithms allow for the restriction in the flow intensity ISDN, working in a variety of conditions and can be used to limit the intensity of flows when accessing the control zone, so; and to limit the intensity of the flow when accessing GAM within the control area. Selection of a particular algorithm IPR defined compromise between implementation complexity and efficiency , which in turn depends on the nature of the internal and mixed-zone traffic parameters IHT and GAM , dimension , zone control , and other factors , which may take into account when using a simulation model of ISDN. Numerical implementation of the algorithm is demonstrated by a specific example.

5. Conclusion

Solving the problem of accurately calculate network performance leads to a significant reduction in mean - time delivery of data packets to a user and reduce loss in the calls. Thus, due to the effective bandwidth redistribution integral paths network can achieve the optimal parameter values in customer service ISDN. All these above factors associated with the problem of optimal allocation of bandwidth between subnets CC and CP, in the end result, significantly increase the effectiveness of multi- network communication in general and thereby actually a transition to a more cost-effective networks.

6. References

[1] Building a Reliable and Secure Multiservice Network for Homeland Security and Defense, Cisco Systems, Inc, 1992-2003.

[2] 15th Progress Report on the Single European Electronic Communications Market – 2009

[3] Tulemisova, G.E. Some mathematical representations of Erlang Formula / Tulemisova G.E. // “Izvestiya vuzov” magazine, #5. – Bishkek: “Scientific magazine and children literature”. 2012. Pages 3-5.

[4] Tulemisova, G.E. A mathematical model for calculating the parameters of the asynchronous data network with circuitous directions / Tulemisova G.E. Izvestiya vuzov” magazine, #5. – Bishkek: “Scientific magazine and children literature”. 2012. Pages 14-18.

[5] Tulemisova, G.E. A method of optimal allocation of channel resources in multiservice network. / Tulemisova G.E. // “Science and new technologies” magazine, #1. – Bishkek: “Scientific magazine and children literature”. 2012. Pages 44-47.

[6] Tulemisova, G.E. Algorithm for calculating the parameters of quality of service digital network with integrated services / Tulemisova G.E., Ashigaliyev D.U., Mustafin S.A. // Materials of Science Forum, dedicated to the 75th anniversary of Prof., Academician MAIN Biyashev R.G. – Almaty, 29.03.2013. Pages 38-42

SESSION

POSTERS

Chair(s)

TBA

An Empirical Approach to Identifying the Friendships on Social Network Users

Kuan-Hsi Chen[1], Yuh-Jyh Hu[1], Tyne Liang[1], and Kiwing To[2]
[1]Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan
[2]Institute of Physics, Academia Sinica, Taipei, Taiwan

Abstract—*Effective friendship identification among users is demanded for most social network services. Nevertheless, such task on a sparse social graph is challenging for network service researchers. Other than the friend database used in most previous researches, the presented study concerns user's textual interactions corpora, containing user's posted messages and received responses. From them, multiple types of interaction features are extracted, including user's sentiment, content topics, response and link information. The proposed identification is built on a stacking learning mechanism and is verified using real social network data. Compared to two other supervised learning models, the proposed approach achieves the highest F1-score and the fastest identification speed. Moreover, the experimental results show that the addressed interaction features indeed play positive roles for friendship identification on social network users.*

Keywords: social network, friendship, user interaction, stacked generalization, feature extraction

1. Introduction

Nowadays popular social network services provide people convenient ways to interact with each other at any place and any time. The network analysis and link identification are more widely discussed than ever before [1][2]. The research presented in [3] employed classification-based link identification by extracting friend connections among users. However such approach may fail in identifying the links between those nodes which do not have any common neighbors. In fact, the user textual interactions are informative, characterizing the difference among friends and strangers. As pointed in [4], friends tend to respond shorter messages and contain more sentiment content in their interactions. It is also observed from the collected corpus, the response time among friends is generally shorter than the time among the users who are not friends to each other. In this paper, the friendship identification is applied to the users of Plurk, a popular microblogging service in Taiwan. Both the textual interaction corpus and friendship network are employed for feature extraction. In total, twenty features are extracted, including the presented twelve interaction features and eight friend features described in [3]. The identification is implemented by a stacking algorithm and is compared with Naïve Bayes and SVM (support vector machine). The Experimental results show that our method yields the best F1-score and the fastest identification.

2. The Methodology

2.1 The Data

The data used for system development and testing were acquired from 5,628 Plurk users and their friends. The data contain 13,185,045 unique friend pairs, 13,271,338 messages and 83,839,889 responses all posted during August 1, 2011 to February 16, 2012. The training data include 20,000 friend pairs randomly selected from the friend pair lists and 20,000 non-friend pairs which are randomly coupled from 5,628 users and are not in the friend pairs. There are five sets of testing data, each containing 4,000 friend pairs only and are set aside from the training data.

2.2 The Method

The collected Plurk messages and responses are preprocessed by our Chinese textual processing tools. We refine the interaction features addressed in [3] and extract additional ones such as response time and word usage from the user interactions. Table 1 describes all the features used by the stacking-based identification and the compared models.

Stacking has been proved to be successful for many applications in information extractions since it can combine the results of different generalizers and minimize error rate. It is implemented in two steps. In level-1, five logistic functions (equation 1), each corresponding to one feature type, are trained. The outputs of the level-1 logistic functions are taken as points of a new feature space for the level-2 logistic function. The level-2 function is also trained by equation 1 to identify the friendships among users.

$$f(X) = \frac{1}{1 + e^{-\sum \theta_i x_i}} \quad (1)$$

3. Experiments

The identification performance is evaluated in terms of precision, recall and F1-score. For model comparison, the supervised methods, SVM and Naïve Bayes, are also implemented and tested in the same experimental settings. Table 2 shows that the stacking method yields the best F1-score by modeling the effect of each feature type independently and

Table 1: The features.

Type	Feature Description (note: user u_i)
Link	the messages to which both u_i and u_j respond
	the authors to whom both u_i and u_j respond to their messages
Sentiment	the degree of sentiment agreement expressed in the Plurk threads that both u_i and u_j participate
	the degree of sentiment disagreement expressed in the Plurk threads that both u_i and u_j participate
	positive emoticons in responses between u_i and u_j
	negative emoticons in responses between u_i and u_j
Content	response length of u_i and u_j responses
	URL sharing in u_i's and u_j's responses
	word overlap of nouns and verbs of u_i's and u_j's messages and responses
Time	average responding time between u_i and u_j
	response count of the most recent month between u_i and u_j
	time gap of the most recent message and response
Friend	common neighbor
	Admic/Adar Index
	Jaccard Index
	preferential Index
	u_i's friend count
	u_j's friend count
	opinion leader
	social balance status

Table 2: Identification results.

No.	Classifier	Feature	Precision	Recall	F1-score
1	Naïve Bayes	Friend	97.5	75.2	84.9
2	Naïve Bayes	All	98.3	82.5	89.7
3	SVM	Friend	86.0	92.7	89.2
4	SVM	All	96.0	92.0	94.0
5	Stacking	Friend	95.2	90.5	92.8
6	Stacking	All	95.9	94.4	95.1

the outputs of level-1 logistic functions of all feature types are maximized so that better results are generated. Table 2 also shows that all the methods obtain better results when all features are used. Moreover, the stacking method using all features (indicated as classifier #6 in Table 2) identified additional 1,075 friend pairs out of 20,000 pairs than the one without interaction features (indicated as classifier #5).

As shown in Fig. 1, the proposed interaction features are effective in identifying those friend pairs containing few mutual friends. Here, we categorize friend pairs into different groups in accordance to their mutual friend count and calculate the cumulative F1-scores. All the methods using friend features only (line 1, 3, and 5 of Fig. 1) are limited by mutual friend information.

Moreover, the proposed stacking based model (implemented in C language) is 100 times faster than SVM (implemented by LibSVM). We believe that it is easier to deploy the stacking based friendship identification method on a distributed system since the outputs of level-1 base functions can be calculated separately.

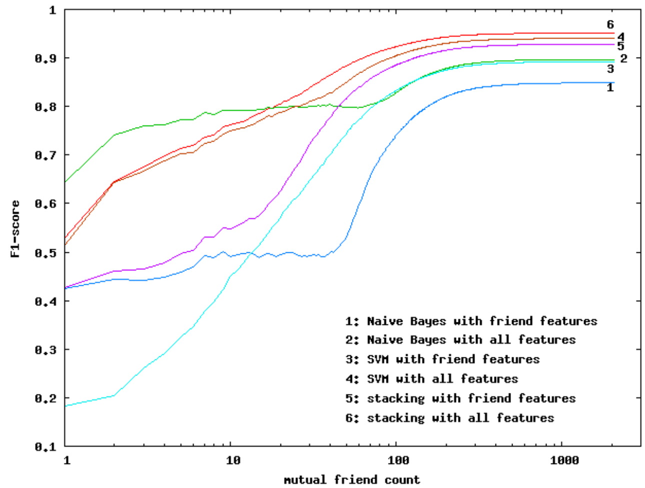


Fig. 1: Cumulative F1-score

4. Conclusion

Undoubtedly, an effective friendship identification facilitates people's interactions through social network platforms. In this paper, we present a stacking-based identification using multiple types of the features extracted from both user interactions and user profiles. Compared to Naïve Bayes and SVM approaches, the proposed identification achieves the highest F1-score and fastest identification speed. In addition, the addressed user interaction features are proved to be complementary for the friendship identification on a sparse social network graph.

Acknowledgement

This work was supported by National Science Council of the Republic of China under Grant No. NSC 101-2221-E-009-148-MY2.

References

[1] D. Liben-Nowell and J. Kleinberg, "The link prediction problem for social networks," in *Proceedings of the twelfth international conference on Information and knowledge management*, ser. CIKM '03. New York, NY, USA: ACM, 2003, pp. 556–559.

[2] S.-H. Yang, B. Long, A. Smola, N. Sadagopan, Z. Zheng, and H. Zha, "Like like alike: joint friendship and interest propagation in social networks," in *Proceedings of the 20th international conference on World wide web*, ser. WWW '11. New York, NY, USA: ACM, 2011, pp. 537–546.

[3] Y. Dong, J. Tang, S. Wu, J. Tian, N. Chawla, J. Rao, and H. Cao, "Link prediction and recommendation across heterogeneous social networks," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, Dec 2012, pp. 181–190.

[4] K.-H. Chen and T. Liang, "Friendship prediction on social network users," in *Social Computing (SocialCom), 2013 International Conference on*, Sept 2013, pp. 379–384.

SESSION

LATE BREAKING PAPERS AND POSITION PAPERS: INTERNET COMPUTING AND BIG DATA

Chair(s)

Prof. Hamid R. Arabnia

A Real-time, Cloud-based Architecture for Integrated Intelligence Driven Operations

Jeffrey Wallace, Ph.D.
I2 Sentinel Associates, Inc.
Fairfax, VA 22030
wallacej@i2sentinel.com

Professor Thomas Preston
George Mason University
Fairfax, VA 22030
prestoth@nro.mil

***Abstract*— Many "big data" software systems are not interactive, automated, or run in a real-time mode. The true utility of cloud computing and "big data systems" can be increased by providing an execution framework and control software that is native to cloud architectures and supports interactivity and time synchronization. In addition, a framework to integrate different artificial intelligence and machine learning algorithms is combined with the execution framework to create a powerful cloud computing system development platform.**

Keywords—cloud computing; analytics; complex system representation; integration; interoperability; conceptual graphs; prediction; intelligent systems; anticipatory intelligence; timing; synchronization

1. INTRODUCTION

Current cloud computing application development and systems integration suffer from the following problems:

- ***Batch oriented*** – the systems are not made to run interactively, or continuously.
- ***Brittle*** – easily fail with slight perturbations to the information transacted
- ***Difficult to maintain*** as systems are upgraded
- ***Complicated to scale*** when adding additional information and constituent systems
- ***Cannot provide clear, concise implementations*** among diverse sub-system sets

Developing operationally relevant computational capabilities requires access to a more cost-effective, compute-focused path. Some jobs or tasks require high performance computing platforms and some do not – the goal is to optimize execution based on available assets. This will be accomplished using an orchestrating system leveraging the runtime performance, throughput and features of enterprise supercomputing platforms. Additionally, this architecture and its realization employ highly optimized software coding, algorithms and applications, specific hardware, benchmarks and logic workflows to achieve the premier supercomputing conditions.

Consolidation and execution of high performance and supercomputing (HPC-SC) systems – the goal is to provide a heterogeneous environment where resource and task are optimally matched. Understanding the interplay between algorithm and hardware is key to optimization at the architectural level.

A cloud application development and integration framework for intelligent systems, called the Joint Execution Environment (JEE) [1], address these challenges. JEE's biologically inspired characteristics have the ability to:

- Create complex, realistic, and scalable networks of component inter-relationships
- Distribute autonomous controls and monitors
- Implement complex webs of cause and effect
- Dynamically alter the execution structure
- Adapt and evolve the system.

The goal is to build systems where decision algorithms are integrated and interoperable with a wide variety of system components. Such "intelligent algorithms" can now be capable of implementing robust monitoring and active control systems that have human oversight and intervention. The speed of the battlespace required that systems ***anticipate*** and handle that gap in time when data and measurements exist to draw a conclusion with a high probability that something is going to happen, and any human decision-maker drawing the same conclusion.

Systems, responses, and countermeasures can now be activated in advance of events that negatively affect operations. Consider the task of controlling a ground vehicle in an urban environment. Standard "rules of the road" provide the example for one type of behavior – that of following either crisp of fuzzy rules via an inference or rules engine. Now consider the case where children are present. The system must recognize the presence of children, which is most efficiently calculated employing a fundamentally different algorithm type than a rule or inference system.

To implement a robust system response, or ***anticipatory intelligence***, advancement beyond traditional Bayesian networks, sometimes referred to as belief networks, is required. A Bayesian network[1,2], roughly speaking, would say that since the probability is higher of something suddenly coming into the vehicle path, given the presence of children, speed should be lowered and a higher scan rate/decision cycle until the hazard is passed.

As stated, entity extraction from unstructured data, or recognizing something from an image of video stream, is a set of algorithms and processes that must be coordinated in the context of other decision algorithms and processes, such as feeding the Bayesian network monitor for road hazards. The goal is to integrate multiple algorithms as appropriate to provide robust automated systems with anticipatory intelligence. What is required is an approach that naturally permits algorithm integration, in addition to system component integration and interoperability.

In addition, JEE addresses the challenges by fusing:

- Advanced systems theory and practice
- Advanced software development
- Low-latency, high throughput, reliable, and robust computer communications
- Sophisticated software integration, interoperability, and synchronization

Common approaches to complex system infrastructure, such as systems based on Microsoft's .NET framework [5], process-based programming (e.g., systems utilizing threads, semaphores, and locks) [6][7][8], object request brokers [9][10], ERP infrastructure [11][12][13], and cluttered web-based technologies, [16] fail in one or more of the problem areas listed above. The tremendous number of constructs causes significant setbacks with most application development and integration methods. JEE (with a macro-based sub-language) easily represents and constructs these complex system capabilities.

As outlined by the Chairman of the Joints Chiefs of Staff [2], it becomes imperative to support mission command by effectively leveraging, and/or developing and distributing the requisite technological solutions to align with situational and regional needs. Reference Section 6, Sub-Factor 4: Innovation and Vision, Paragraph 1.

2. MULTI-INT ARCHITECTURE

This architecture enables intelligence community (IC) and Department of Defense (DoD) standards including the use of the IC Trusted Data Format (TFD)

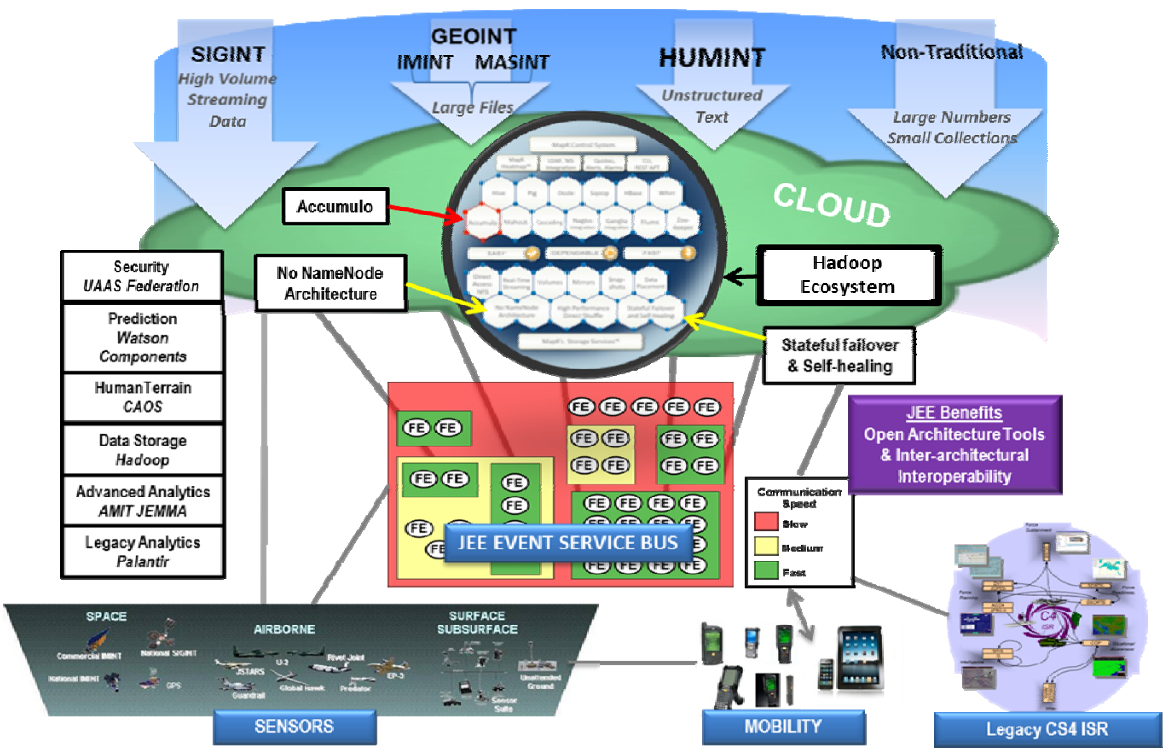


Figure 1 Real-Time, Multi-INT Architecture. *The architecture solution improves information content and flow to the warfighter using superior regional domain awareness through Joint Execution Engine (JEE).*

1 http://en.wikipedia.org/wiki/Bayesian_network

2 http://www.dmoz.org/Computers/Artificial_Intelligence/Belief_Networks/Software/

and smart data tagging schemas including the Enterprise Data Header (EDH), Access Rights Header (IC-ARH) and Information Resource Metadata (IC-IRM) as components of the Standard Data Representation 2.0.

Appropriate interfaces to the IC's Unified Authorization and Attribute Service (UAAS) federated Identity and Access Management (IdAM) service (ICITE/ AA provider) and the community cloud, widget, widget interface and data interface standards are available.

Collaborative integrated intelligence efforts within the IC (notably object based production(OBP), Red Force Tracker (RFT) pilot) to facilitate unifying information, conforming to the IC in shaping and organizing the "known," as Object Based Targeting (OBT) expands and scales to new intelligence domains (called INTs), and object types will increase the effectiveness of monitoring the areas of responsibilities (AORs). Our approach implements Object Management as a Service (OMaaS) to enable users, through war gaming and discovery for governance, policy, security and accountability processes associated with OBP.

In particular, intelligence can be linked to known Objects facilitating semantic methods and understanding. This approach will expand OBP capability to the any AOR, improving situational awareness through large object identification to include named areas of interest, missile and launch facilities. Additional INTs may include near real-time video, infrared and MASINT to improve situational awareness. The objective is to establish end-to-end performance of calculating new knowledge and determine how to send new information to the Global Command and Control System (GCCS) for improved understanding.

2.1 Real-time Access and Understanding of Multi-agency Data

Real-time access and understanding of intelligence data is the hallmark of this approach. The infrastructure has been architected to enable true real-time system integration, interoperability, and implementation. Every algorithm has been extensively analyzed and its implementation optimized.

A key differentiator to existing enterprise offerings is the ability of this infrastructure to be ported to any operating system and compiler set. This enables direct integration between IT, C4ISR, weapons systems, and emerging unmanned systems in a way previously not possible.

Several technologies are combined to implement real-time access and multi-agency understanding. A key implementation feature is the incorporation of Accumulo product within an enterprise hardened commercial version of the Hadoop technology ecosystem. Our approach unifies the current efforts by various agencies and leverages commercial versions of technologies that include advanced enterprise features focused around performance, reliability and availability. The improvement in this version of Hadoop is the elimination of a single name node. We eliminate a key bottleneck in the existing architecture and enable a system that delivers linear performance as the size of the system grows.

Our implementation also features an expanded set of ingest APIs and methods that permit real-time streaming and other standardized file and database system interfaces. These factors improve performance, integration, reuse, and interoperability.

Expanded interfaces also permit the usage of a diverse commercial analytics, data, analysis, and data management tools. Our technical approach also permits the usage of high performance solid state storage technology that can revolutionize the manipulation and analysis of large objects and real-time, or near real-time, performance of data analytics. A HPC-SC event service bus is employed to enable data analytics scheme to be implemented. The event bus infrastructure enables programming complex orchestration while delivering real-time performance. The nature of the bus permits inter-architecture integration between multiple systems simultaneously.

2.2 Advanced Analytics Capability to leverage High Performance Environment

Tailored algorithm implementation is the key to optimizing advanced analytics for HPC-SCs to understand how to phrase the analytic algorithms to avoid processing bottlenecks. Algorithm optimization occurs in several different parts of the architecture beginning with the HPC-SC event service bus. The implementation enables event queue distribution over the processing nodes and coordinates an advanced synchronization algorithm. We support data marshaling between the data sources and stores in the analytic algorithms and integration and interoperability between analytic algorithms and entire analytic architectures.

This technical implementation permits the federation of heterogeneous analytic architectures in real-time providing not only improved performance but the ability to utilize all interagency and operational analytics.

A wide variety of libraries of algorithms, data structures, and message distribution and synchronization constructs which are optimized for HPC-SC infrastructure will be provided. This enables analytics to be parallelized leading to drastically reduced run times. Analysis performance can also be enhanced by the inclusion of graphics processing engine providing the system with the ability to support emerging GPU APIs that can support the type of vector and matrix operations required by advanced analytics. Such a polymorphic computing architecture optimizing algorithm implementation and execution, in addition to algorithm scheduling and synchronization, provides a comprehensive approach to the full utilization of HPC-SC resources.

2.3 Command Support

Central to the proposed architecture and implementation is the ability to integrate and interoperate with existing military command information, C2, and C4ISR systems. This is accomplished thru the external module interface framework in the overall HPC-SC event service bus framework, Object Based Production, Red Force Tracker, Army Advanced Analytics, GCCS I3 COP, NGA GEOINT, and other OZONE Widget capabilities.

This implementation includes a gateway for GCCS, Army Battle Command System (ABCS) publish and subscribe system, Command and Control PC (C2PC), Command Post of the Future (CPOF) and the Theater Battle Management System (TBMCS). This will provide a framework and re-useable code base intended to incorporate other sensor interfaces and systems as necessary for instance, integrating the sensor suite from the Global Hawk/Trident surveillance programs.

The ability to integrate new analytics and permit composition to meet emerging requirements is fully optimized and supported by the framework. For example, systems multiple domain analytic systems will be integrated and incorporated easily. The framework has a natural spot for the inclusion of security and the ability to support cross-domain solutions. This approach will also revolutionize the use of personnel assets in terms of their location in the Battlespace / AOR.

2.4 Advanced Analytical Capabilities, Framework, Data Layer, and High Performance Computing System

This approach permits the integration or implementation of any new analytic tool or algorithm. This effort will provide a framework for the ingestion of all source data, both structured and unstructured; expand the heuristics for problem solving, learning and discovery, and provide mediation and normalization unifying access to and analysis of all data. Specifically this effort provides several new commercial analytics tools and utilities.

A novel video and sensor analysis classification and metadata tagging tool and framework will be provided so that for the first time, full motion video (FMV) and other sensors can be classified and metadata tagged upon ingest, and both the object and the metadata can be streamed into a unified Data Layer. This increases the value of the data for a wide variety of analytics and provides critical improvements to intelligence and operational decision systems and processes, leading to faster production of decision quality data and acceleration of the mission command and the decision process.

A variety of analytics engines are featured that will be employed capable of creating indexing schemes which can be used separately or combined in the decision process. This capability will lead to the advancement of analysis, intelligence and decision processes in new ways. For example an index can be created based upon the ingestion of a variety of classified data sources and integrated with indexes from other departments or classification levels without compromising source data or methods and improving analytics queries.

Combining capabilities together is the HPC-SC event service bus and associated HPC-SC optimized support libraries, packages, and utilities. This team will work with both commercial vendors and relevant Government and academic analytics packages and technologies to improve their native utilization of HPC-SC technologies and development methodologies to enable individual analytic tool performance.

2.5 Trusted Systems

One of the true challenges is the computationally efficient trusted systems. The need to constantly enforce security policy at every stage of operation is a significant overhead. The JEE provides a set of execution control primitives that are “lightweight” from the perspective of memory and processor cycle consumption.

The development of trusted systems is clearly associated with three driving development areas:

1. Enterprise testing
    a. Create the necessary baseline
    b. Freeze the baseline and testing
    c. Maintain the state
2. Representing composite relationships within the cloud
    a. Functions and functionality
    b. Identify where the work is being performed
    c. Understand how we establish trust in the cloud
3. Implement trust in the cloud
    a. Emphasize workload and metrics
    b. Constant software modeling, simulation, and coding from test to operational readiness
    c. Code management
    d. Repeatable, auditable, and demonstrable

This architecture and implementation demonstrates how cloud and system tuning can be simplified by allowing all aspects of the hardware and software infrastructure to be optimally accessed. State-of-the-practice enterprise systems always want to define their own “pure” architecture based on some seemingly relevant philosophical concept. This approach allowing the most efficient system functions, and when trust is implemented the systems perform poorly.

Regarding how cloud computing techniques and technologies could be used to enhance interoperability and improve the sharing of geospatial information across a broad range of users and devices, the architecture allows both high performance and resilience to system failure. Geospatial information sharing in operational environments must be resilient to infrastructure attacks and failure. This approach allows not only highly available “traditional” resources in terms of compute and data storage, but also in terms of communication. In all other systems, a communication link problem results in data transmission failure, whereas this architecture and implementation feature the ability to have completely reliable, yet efficient, data communication.

This approach provides a consistent and secure single map solution for appropriate users through open APIs. A wide variety of messages, formats, and protocols can be used based on what different communities indigenously support. This avoids costly enterprise standardization, and focuses on letting communities immediately interoperate, while providing a smooth evolution path to increased shared development operations (DEVOPS), resulting in a natural technical convergence. Legislated, top—down convergence efforts rarely succeed, and when they do it is only at great cost.

This architecture also enable success in the development of broad international collaboration efforts by virtue of the flexible nature of data sharing APIs and methods. Many partner nations utilize a large quantity of legacy systems, and modern enterprise approaches do not easily support older systems and alternative architectures.

3. ARCHITECTURE IMPLEMENTATION

The architecture shown in Figure 1 will improve information content and flow to the warfighter, enabling superior regional domain awareness through a robust synchronization capability inherent in the Joint Execution Engine (JEE) software component. This is provided as an optimizing and enhancement mechanism to the HPC-SC and Enterprise Service Bus (ESB) functionality currently used in the IC. The JEE introduces a mediation layer that effectively enables introduction of analytics to Map Reduce in the cloud. This also introduces a multi-source input, a co-creative environment and other commercially developed efforts in guided information search/discovery/filtering to optimize information flow to the warfighter.

Data flow modeling is employed to document flow through the interface to ensure data credibility from source to product. Through this interface we will ingest C4ISR resources to support strategic and tactical data into the data layer to include emerging advanced analytics widgets. Our proposal highlights innovation in several dimensions that address requirements as set in the above. The design and implementation of the Data Layer delivered via the JEE and its integration of key products, unifies the inter-agency Data Layer strategies.

The Data Layer ingest is significantly expanded to permit streaming data, POSIX file system files, and existing database standard interfaces. The implementation provides an enterprise level by the automation of mirroring and other persistence services in the Data Layer. The architecture supports the existing analytics and Widget framework.

The scope of real-time support is extended by employing a high performance, real-time event engine optimized to provide real-time, continuous analytics support to the Warfighter. Existing C2 and C4ISR systems and envisioned future Warfighting capabilities including unmanned systems, future sensors, automation frameworks and intelligence sources are supported. The HPC-enabled event bus permits the orchestration of widgets, Map Reduce jobs, any commercial analytics, other analytics models (i.e., SQL, POSIX, or JDBC/ODBC), and simulations as desired by an AOR or IC user.

3.1 JEE Architecture Description

JEE is an object-oriented, event-based, high performance execution system. JEE [1] provides high speed communications, which is central to its framework [2], and utilizes numerous messaging fabrics for inter-processor communication: shared memory, wireless, fiber optic, ATM, TCP, IP, and multicast (implemented in a variety of media). The **Communications Services** provide a variety of mechanisms linking clients to intelligent application services, or the hosting processors. **Communications Services API** (internal and external) is standardized for simplified integration [1]. An abstraction is then supported for unicast and multicast, and permits various implementations, and thus protocols, to work simultaneously.

State-Saving Framework and State-Saving Services support reliability, synchronization, fault-tolerance, and implementation of persistence services. The Core Programming Services provides Standard Template Library (STL) programming API, and utilizes the state-saving and persistence features.

Event Management Services (EMS) provide high-performance data structures to develop exceedingly complex, reliable, interactive intelligent applications more rapidly.

Synchronization Management Services, closely coupled with EMS, control synchronization and timing, which are essential for real-time intelligent applications interfacing with hardware [2].

Standard Application and Integration API (SAIA) expedites development of complex, robust intelligent applications by code generation macros and APIs. SAIS synchronizes components as the overall system executes, and scales to simultaneously execute large numbers of components (an interactive synchronization mechanism hides programming complexity).

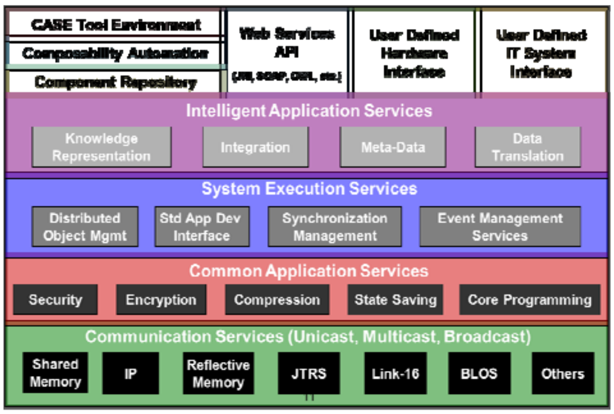


Figure 2. JEE architecture layering

Distributed Object Management Services and **Data Translation Services** provide location transparency and a powerful, yet easy to use, distributed object computing framework (e.g., the complexity in using different inter-processor communications).

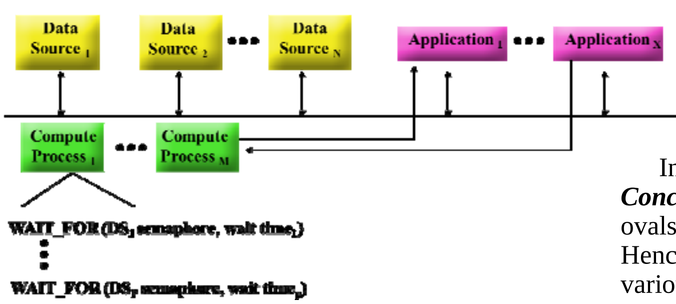


Figure 3. Sample JEE System Layout

The **Metadata Infrastructure (MDI)** provides:

- A syntactic notation for relational associations
- Logical constraints, consequences, and role-based behaviors entailed by the relationships
- A system of primitive relational abstractions
- Method to change abstractions into conceptual hierarchies
- Means to compose and structure relationships into frameworks and architectures

JEE utilizes additional state of the art technologies in its architecture: **Knowledge Execution Engine (KEE)** and **Knowledge Representation Integration Infrastructure (KRII)** employ the lower layers for location transparency, and are based on ***conceptual graphs*** **(CGs)** [3].

Collectively, the extensions embodied in JEE signify precision in the art. All services adhere to the **External Integration Framework**, and easily integrate intelligent applications and WAN (Figure 2) [1].

4. INTELLIGENT SYSTEMS

The JEE utilizes CGs [1] to integrate hardware and software systems. Visually, a CG mimics knowledge representation in common diagrams for discussions (using whiteboards, slides, or even table napkins). The diagrams, or drawings, are often text snippets (typically enclosed in squares or ovals) and lines (such as labels) connecting one snippet to another. Experts often use visual aids to quickly communicate complex details during brainstorming sessions (see Figure 4).

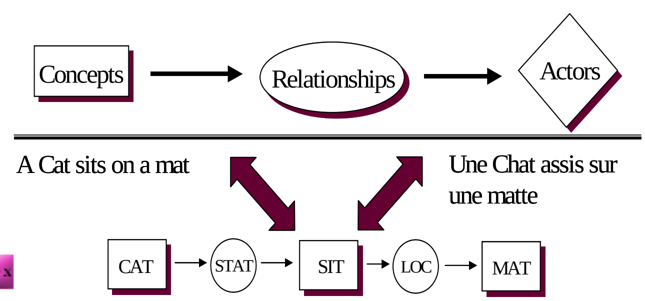


Figure 4. Basic CG example

In CGs, text snippets (in a square) are called ***Concepts***. The line connections are enhanced with ovals**,** called ***Relationships***, containing additional text. Hence, representation of semantic relations, between various concepts, occurs in a manner consistent with common "brainstorming pictures". ***Actors*** (diamond shaped symbols) provide a method to encapsulate interfaces to hardware or software components, and indicate data or signal transforming activity is occurring.

Structurally, a CG provides the following advantages for representing and integrating complex systems:

- Inherently hierarchical; permits operation at increased aggregated levels, when beneficial
- Can decompose components to appropriate levels of detail, to meet requirements
- Ability to conceptualize the entire system, or one specific concept
- Capable of capturing any aspect of the system

Using CG's, JEE simplifies hardware and software component integration; and simultaneously, concisely represents the entire system control logic.

Developers typically manage functional blocks with various data sending and receiving protocols. However, most developers lack standard approaches across the enterprises (usually one method works as well as the next). Consequently, accurate prediction of time and cost is difficult. JEE provides an organized template for encapsulating functional blocks (hardware or software), which is easily mastered by integrator engineering staffs, thereby solving the prediction challenge.

CG Relationships have simple rules - one concept node must be connected by an incoming arc to a relationship; and one concept node must be connected by an outgoing arc. Relationship nodes provide critical semantic structure to system descriptions, and frequently represent modifiers, qualifiers, and constraints.

Concept nodes may be connected to relationship nodes or actor nodes, but direct connections between concept nodes is not permitted. A concept node may have any number of incoming arcs or outgoing arcs, and represent a variety of system features. Concepts may be components (or objects), or they may represent actions (or verbs). Relationships act as modifiers (adjectives and adverbs).

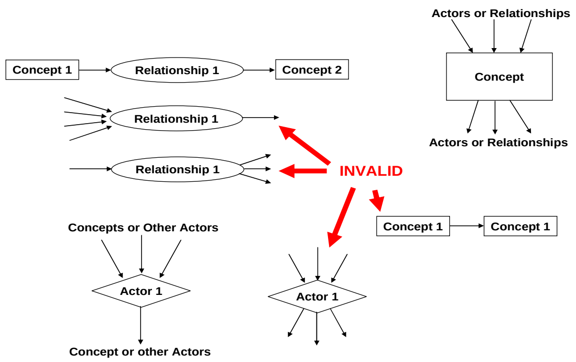


Figure 4. Conceptual Graph creation rules

Actor nodes can have any number of incoming arcs from other nodes. However, they have only one outgoing arc. The outgoing arc may be connected to another actor node, or a concept node. Actor nodes provide the critical ability to encapsulate hardware or software components. Hence, actor nodes enable system integration.

One of the purposes of the Knowledge Execution Engine (KEE) involves controlling the execution of a collection of hardware and software components, as a cohesive and robust system. Relationships have the simplest rules - one concept node must be connected by an incoming arc to a relationship; and one concept node must be connected by an outgoing arc. Relationship nodes provide critical semantic structure to system descriptions. They frequently represent modifiers, qualifiers, and constraints.

5. IMPLEMENTATION EXAMPLES

One of the central tenets of the JEE is to make the programmers job simpler by virtue of API design. The act of sending data from one component to the other can often be a challenging task in modern enterprise implementations. Monolithic, complex data models must be maintained, and complied with, leading to what practically results in an expensive and slow moving social engineering process.

In addition, the state-of-the-practice for the implementation of data interoperability has resulted in a philosophy that requires internal data structures to be rectified and flattened into text files in XML syntax. Ideally, the natural data structures used by developers would be sent, and there would be no monolithic data model. Any one component typically only required the interaction with a small set of other components, so practically speaking this should be decoupled technically and in process implementations.

5.1 System Synchronization

One of the most difficult tasks is the precision implementation of causality between system components. The current state-of-the-practice in real-time systems focuses on the accurate control and estimation of the time to complete a function. The JEE provides such timing constructs, and extends the paradigm to integrate logical controls with timing. This permits the notion of an interrupt to a timed component, a strictly logical synchronization of components in a real-time environment, and component synchronization timeouts.

Consider the following example where the firing of a gun is controlled over the network or locally. In some fashion the gun engages a target by moving the barrel into position and then firing the weapon. The process of moving a gun into position can be expressed and implemented with **slew** and **elevate** functions, as indicated in the figure below.

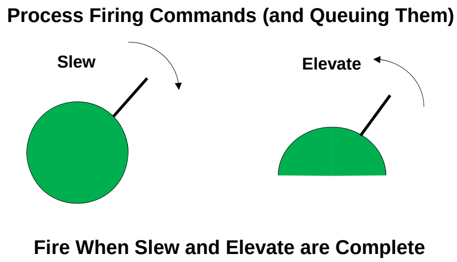


Figure 5. Gun dynamics

The code fragment to implement this is:

```
void Turret::fire()
{
  P_VAR

  P_BEGIN(2)

    // Wait until the turret movement is completed

    WAIT_FOR(1, slewComplete, -1);
    WAIT_FOR(2, elevateComplete, -1);

    // Fire the weapon, this would activate the real gun

    Fire_M256();
    RB_cout << "Flash, Boom, Bang, Echo" << endl;
    fireComplete = 1;
  P_END
}
```

The key to ease of programming is illustrated by the use of programming macros to generate code patterns and implement transparent semantics. The WAIT_FOR macro implements a thread-like suspend and resume, but utilizes the local stack defined in the P_VAR-P_BEGIN block, which drastically reduces memory from megabytes to store the stack frames normally used in thread packages to do the same function to bytes.

The **slewComplete** and **elevateComplete** are types of semaphores, and can be released by network or local API calls. This implements net-centric systems optimally, and with precision. These same primitives can be used to implement trust policies that control execution paths.

5.2 Data Distribution

The following example is from the Joint Strike Fighter (JSF) Shared Synthetic Environment (SSE) Risk Reduction Activity (RRA). In this case, JSF (F-35) data must be sent to a variety of clients. One can envision a JEE collective as depicted below. Some numbers of different types of functional elements (FE) perform computation, and some variable speed communications fabrics provide communication between them. Systems and data sources are plugged in to the collective via an easy to program publish and subscribe API.

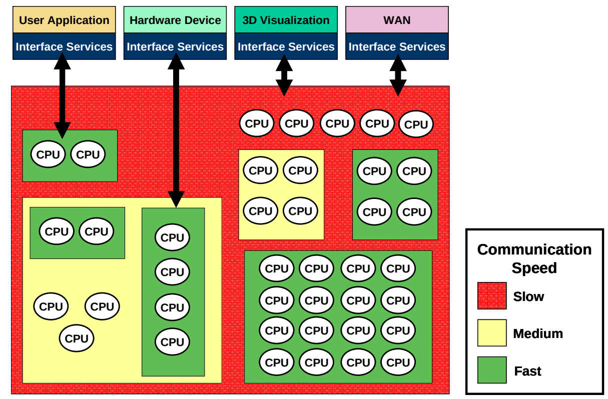


Figure 6. JEE polymorphic computing architecture

The native data structure **StateUpdate** is sent as a segment of binary data. The **CdParameterSet** has the ability to contain both structured and unstructured data in a single package using a series of **Set** functions.

```
platState = new CdParameterSet;
platState->SetBuffer((char*)"UpdateStateMsg", (char *)&StateUpdate, sizeof(UpdateStateMsg));
SCHEDULE_INTERACTION(CdGetSimTime(), (char*)"Update.Ownship.PVT", *platState);
```

On the receiving side in a system to be integrated with the JSF would insert the following code in order to operate with the JSF state data.

```
CdEmf *Emf = NULL;
...
...
Emf->SubscribeInteraction((char*)"Update.Ownship", &OwnshipUpdateHandler, 1);
...
...
void OwnshipUpdateHandler(CdSimTime time, char *interactionName,
                CdParameterSet &interactionParameters) {

  char *StateMsgBuffer;
  UpdateStateMsg *State;
  int numBytes;
// cout<<"In Ownship Handler"<<endl;
  interactionParameters.GetBuffer((char*)"UpdateStateMsg", StateMsgBuffer, numBytes);
  State = (UpdateStateMsg *)StateMsgBuffer;
...
}
```

The ability to simply send any collection of data in a lossless manner is now possible. The developer needs to know nothing of the communications method, security method or controls, encryption, or compression being used. The code is then more resilient and flexible at the developer level. This approach permits code in Java, Python, Ruby, or PHP.

5.3 System Control

Consider a more complex scenario than driving, such as large-scale battlespace threat monitoring and response, specifically that aspect regarding anti-access/area-denial (A2AD). There is always an order of battle, or sequence of inistializing actions that have to take place before any complex offensive or defense operation, the goal is to have key sensors that can collect, ideally, sufficient statistics for relevant phenomena, infer meaning from those measurements, such as from a Baysian, or belief, network, and act on that information.

The term proposed for this new pattern of activity is collect-correlate-alert-act (C2A2). Various algorithms can now be inserted as part of the mechanism that actually triggers activity. In the case of battlespace monitoring consider the case where a connection is lost with an unmanned aerial vehicle (UAV) and a satellite. An alert is generated warning of a possible electromagnetic pulse attack, and appropriate actions are taken such as powering certain systems down and activating others, such as a mobile AOC.

In other words, based on multiple and certain sequences of events, both systemically and socially there exists a probability of event that would necessitate activating AOC (M) system, thereby placing a pilot (UAV or in-cockpit) in an advantageous rather than disadvantageous situation in a potentially denied area, as opposed to having to improvise independently due to the slow speed of human and organizational coordination. The goal is to not create another Situational Awareness System, but a cognitive/active anticipatory system that reduces and reverses the disadvantage being created by our adversaries.

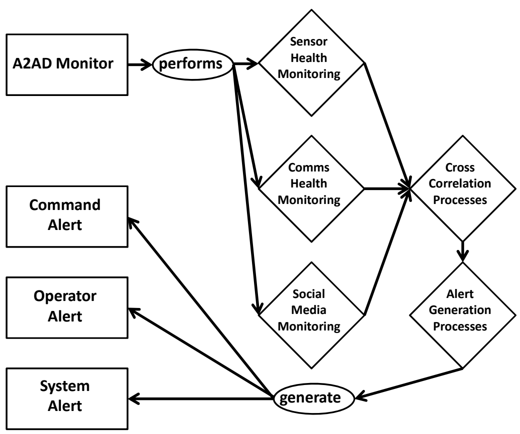


Figure 7. A2AD conceptual graph

```
CdConceptual Graph A2ADMonitor;
…
A2ADMonitor.Initialize(a2ad.owl);
A2ADMonitor.Execute();
…
CdConceptualGraph::Execute()
{
P_VAR
P_BEGIN

while (Active)
  WAIT_FOR(1, systemInterrupt, scanInterval);

P_END
}
```

The code to implement this is straightforward. A conceptual graph entity is initialized with the CG in figure 7, and then periodically executed. The complexity is in the actor subgraphs, and these can be modified by reloading the top level CG, or the CG replaced completely. This allows the system to adapt in terms of behavior, and with dynamic binding we can even add new functional components.

6. SUMMARY

There are many advantages of this approach to complex, automated, intelligent system development. Key advantages include ease of use, with an intuitive implementation framework for developers; and reduction of rework – getting the job done right the first time. The JEE helps meet practical needs in system development. Frequently testing, security, and interoperability are considered too late; and capitalizing on software engineering research rarely occurs.

Furthermore, existing assets can be re-cast and given "a new life"; and existing hardware assets have a longer lifespan. As such, the next generation of solutions is enabled, which mitigates many of the major problems associated with application development and systems integration. Scalable, robust solutions are provided for complex, real-world problems.

7. REFERENCES

[1] Wallace, Jeffrey and Sara Kambouris, 2014. "Intelligent System Development and Integration for Cloud Computing," Proceeding of the 2014 International Conference on Artificial Intelligence, IEEE Computer Society, Las Vegas, NV, March 10-12.

[2] Dempsey, Gen Martin, 2012. "Capstone Concept for Joint Operations: Joint Force 2020," Joint Chiefs of Staff, September 10, 2012.

[3] Peirce, C.S.: Collected Papers of Charles S. Peirce, Hartshorne, Weiss, and Burks (eds.), Harvard University Press, Cambridge, MA, 1932.

[4] Sowa, J.F., Knowledge Representation: Logical, Philosophical, and Computational Foundations. 2000, Pacific Grove, CA: Brooks/Cole, pp. xiv+594.

[5] Microsoft .NET Framework, http://msdn.microsoft.com/en-us/vstudio/aa496123 , Microsoft, Inc., 2014.

[6] Parallel Virtual Machine, http://www.csm.ornl.gov/pvm/pvm_building.html , Oak Ridge National Laboratory, 2005

[7] The ADAPTIVE Communication Environment, http://www.cs.wustl.edu/~schmidt/ACE.html , Washington University, 2011.

[8] POSIX Threads Programming, http://www.llnl.gov/computing/tutorials/pthreads/ , Lawrence Livermore National Laboratory, June 30, 2013.

[9] Real Time CORBA with TAO, http://www.cs.wustl.edu/~schmidt/TAO.html , Washington University, 2013.

[10] IBM Component Connector Overview,http://www.redbooks.ibm.com/abstracts/SG242022.html?Open, IBM, Inc. June 19, 1998.

[11] SAP Solutions, http://www.sap.com/solution.html, SAP AG, 2014.

[12] Oracle Enterprise Resource Planning, http://www.oracle.com/us/products/applications/enterprise-resource-planning/overview/index.html, Oracle, Inc. January 2014.

[13] Oracle Service-Oriented Architecture, http://www.oracle.com/us/products/middleware/soa/overview/index.html, January 2014.

Determining Human-Perceived Level of Safety in Transportation Systems using Big Data Analytics

Carlos E. Otero, Marco Rossi, Adrian Peter and Rana Haber
Florida Institute of Technology, Melbourne Florida, USA
cotero@fit.edu, mrossi2013@my.fit.edu, apeter@fit.edu, rhaber@fit.edu

***Abstract*—Road safety continues to pose a significant problem in the United States. Although significant time, effort, and money have been spent to improve transportation management systems and decrease accidents, statistics show that this epidemic continues to remain unsolved. This highlights the need for advanced data systems that complement existing technology and help identify and solve today's challenging safety problems in transportation system. To this end, this paper presents work-in-progress to create a big data analytics system for transportation safety management. The initial work focuses on determining human-perceived level of safety in areas of emphasis within transportation safety.**

Keywords—Big Data, Internet Computing, Sentiment Analysis, Clustering, Intelligent Transportation Systems

I. INTRODUCTION

Road safety continues to pose a significant problem for the Florida Department of Transportation (FDOT). According to a 2013 report by the National Safety Council (NSC), Florida's rate of car accident deaths is one of the nation's highest [1]. The NCS reports that the number of fatal car accidents in Florida in 2012 was 2,442, revealing a 3 percent increase from 2,372 traffic deaths in 2011. Florida's rate of highway fatalities in 2012 was third-highest in the nation behind Texas, California, and slightly ahead of Pennsylvania. These figures point to inadequacies in Florida's traffic management systems, where enormous human potential is being destroyed, with also grave social and economic consequences. Traffic management that leads to road safety is thus a major public health issue throughout the state of Florida.

Although FDOT spends significant time, effort, and money to improve transportation management systems and decrease accidents, statistics show that this epidemic remains unsolved—not only in Florida, but worldwide. According to the comprehensive World Report on Road Traffic Injury Prevention [2], there is a need for developing *"good data systems for identifying problems and evaluating responses."* Creation of such data systems should help increase *"understanding not only of the volume of traffic deaths, injuries and crashes, but also of which road users are most affected; in which geographic areas the greatest problems are found,"* and *"what risk factors are contributing"*. [2] Although massive and continuous traffic datasets abound in the social media space, no solution exists with demonstrated capabilities to seamlessly integrate this information for creation of such data systems. Harnessing "Big Traffic Data" from social media would provide data systems capable of producing and analyzing data so large that it would contain significant low probability events—something that would be absent from traditional statistical sampling methods. Therefore, there is a need for advancing intelligent transportation systems' technology further into the area of automated massive analytics and visualization of streaming, unstructured data; which complement traditional data collection methods and help solve Florida's traffic management problem.

This paper presents work-in-progress to develop a Big Data analytics system for transportation safety management. The initial focus of the work is on determining human-perceived level of safety based on major emphasis areas identified in the 2012 Florida's Strategic Highway Safety Plan (FSHSP). These include:

- Aggressive Driving
- Distracted Driving
- Intersection Crashes
- Traffic Data

The approach leverages of social media (Twitter) data to continuously monitor transportation safety data; cluster the data into groups representative of emphasis areas, and extract human-perceived level of safety (for particular cluster) in the form of sentiment analysis. The rest of the paper is organized as follows. Section II provides a brief description of the system's physical and logical architectures. Section III describes our clustering algorithm, which is used in the system as a pre-processing step to form groups of similar documents which can be mapped to FDOT's areas of emphasis. Section IV describes the preliminary results of the sentiment analysis algorithm, which is used in this work as main method for determining human-perceived opinion for each document in a given cluster. Finally, section V summarizes the work and provides important areas of future work.

II. System Architecture

The transportation analytic system is modeled using the Unified Modeling Language (UML) to specify a system of systems, where multiple, dispersed, independent systems interoperate to form a larger, more complex system. To support this, the overall architecture is designed using the representational state transfer (REST) architectural style, which (implicitly) includes other important architectural styles and patterns that support service-orientation and distributed computing. The distribution of architectural components is best documented by the client-server architectural pattern, which includes two major components: a server component and client component. To increase flexibility, the server component is deployed as one or more web services, with each service providing its services using a REST-style interface so that clients can initiate requests to servers, servers process these requests, and return appropriate responses. This allows multiple services to be deployed in different locations across geographical sites. The initial system's analytics capabilities include (replaceable) components that provide services for document clustering and sentiment analysis. The client is composed of HTML5 interface that can be used by analysts to access the server's services in distributed fashion

A. The Physical Architecture

The physical view of the system represents the physical or deployment aspects, where the main elements of analysis are nodes, connections between nodes, and the mapping of software artifacts to nodes. Physical machines, devices and processors are reflected as nodes, and the internal construction is depicted by embedding artifacts, which are binary entities depicted by a box with a file icon in the upper-right corner. Artifacts, such as executables, jar files, etc., are allocated to nodes to model the system's run-time configuration. The physical view focuses on modeling of elements that directly affect quality requirements, such as availability, performance, and scalability [3]. In our system, availability is addressed via server redundancy (not depicted for simplicity), which requires downstream design and development to think about techniques for identifying faults and swapping between primary and redundant server nodes when necessary to support the system's availability. Performance is mainly addressed by selection of appropriate server hardware to support throughput and other computational requirements to be specified. A simplified view of the deployment model is presented in Fig. 1.

As seen, the server computer hosts the web services responsible for the overall analytics clustering and sentiment analysis of transportation-related twitter messages. For simplicity, Fig. 1 depicts one analytics service, hosted by one server; however, multiple services can be deployed on multiple computers, collaborating as part of the system. The physical model also shows how binary artifacts are mapped to hardware components, and how logical components are manifested by these artifacts, providing full mapping between the physical (i.e., hardware) and logical (i.e., software) design.

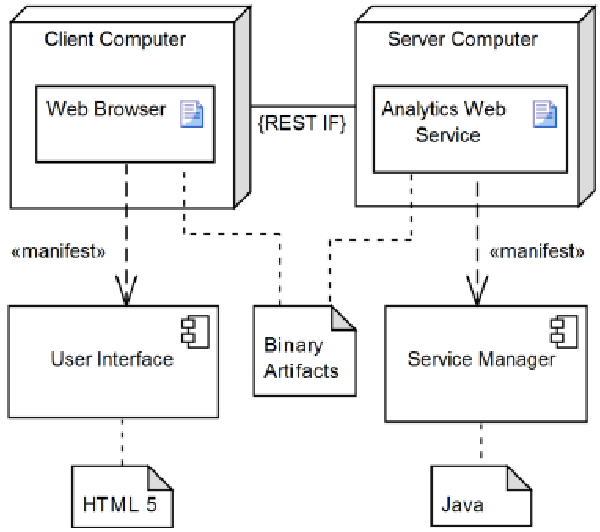


Fig. 1. A portion of the system's physical architecture. Physical devices are reflected as nodes, where artifacts, such as executables, jar files, etc., are allocated to model the system's run-time configuration. The <<manifest>> relationship is used to connect physical and logical architectures.

B. The Logical Architecture

The logical view of the software system is used to decompose systems into logical components that represent the structural integrity that supports functional and nonfunctional requierements.

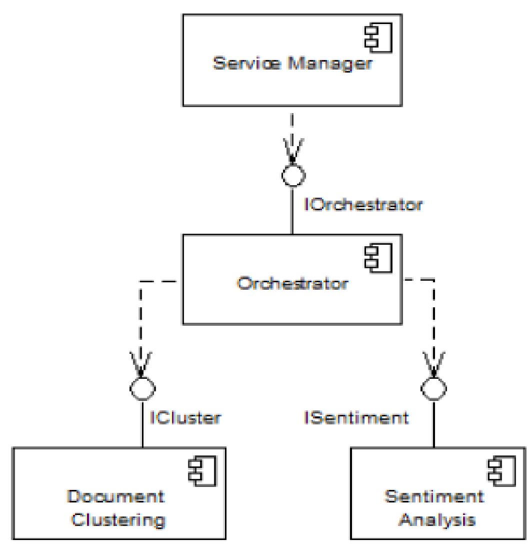


Fig. 2. A portion of the systems' logical architecture. Each component represents an abstraction of a major system function. Components use well-defined interfaces to interface with other components, which makes them replaceable within their environment.

Using this view, the static structure of the system can be modeled using UML component diagrams to decompose, abstract, and encapsulate the services that the system needs to provide to its users. By using these diagrams, the major components, their interfaces, and their associations with all other components are identified. Architectural logical designs exist at a higher level of abstraction and provide the building blocks for further detailed design and construction. They also present the main interfaces between components—these are referred to as stable interfaces, since they expose the main functions provided by each component while hiding the details of their implementation. Stable interfaces are necessary for increasing the extensibility of the system. When using stable interfaces, each component can be replaced with another component that provides similar or enhanced functionality but equal interface. The main logical design for the system is presented in Fig. 2.

III. CLUSTERING COMPONENT

Document clustering is a necessary pre-processing step before sentiment analysis can take place. The ultimate goal is to find clusters of similar messages and relate those clusters to the FSHSP's emphasis areas, including: aggressive driving, distracted driving, intersection crashes, and other traffic data. Many different clustering algorithms exist. An example can be seen using Hierarchical Affinity Propagation. Affinity Propagation (AP) is a clustering algorithm introduced by Frey et al. [4], motivated by the simple fact that given pairwise similarities between all input data, one would like to partition the set so as to maximize the similarity between every data point in a cluster and the cluster's exemplar. An exemplar is an actual data point that has been selected as the cluster center. As we will briefly discuss, these ideas can be effectively represented as an algorithm in a message passing framework. In the landscape of clustering methodologies, which includes such staples as K-means [5], K-medoids [6], and Gaussian Mixture Models [7], predominantly all methods require the user to input the desired number of cluster centers. AP avoids this artificial segmentation by allowing the data points to communicate accumulated evidence encoded in their similarity values, which organically gives rise to a partitioning of the data. In many applications, like document clustering, an exemplar-based clustering technique gives each cluster a more representative and meaningful prototype for the center, versus a fabricated mean.

Hierarchical Affinity Propagation (HAP) introduced by [8] extends AP to allow tiered clustering of the data. The algorithm starts by assuming that all data points are possible exemplars. Each data point is viewed as a node in a network connected to other nodes by arcs such that the weight of the arcs s_{ij} describes how similar the data point with index i is to the data point with index j. HAP takes as input this similarity matrix where the entries are the negative real valued weights of the arcs. Having the similarity matrix as the main input versus the data patterns themselves provides an additional layer of abstraction—one that allows seamless application of the same clustering algorithm regardless of the data modality (e.g. text, images, general features, etc.). The similarity can be designed to be a true metric on the feature space of interest or a more general non-metric [4]. The negative of the squared Euclidean distance can be used as a metric for the similarities. The diagonal values of the similarity matrix,s_{ij}, are referred to as the "preferences" which specify how much a data point j wants to be an exemplar. Since the similarity matrix entries are all negative values, $-\infty<s_{ij}\leq 0$, $s_{jj}=0$ implies data point j has high preference of being an exemplar and $s_{jj}\approx-\infty$ implies it has very low preference. In some cases, as in [4] [8] [9], the preference values are set using some prior knowledge, for example uniformly setting them to the average of the maximum and minimum values of s_{ij}, or by setting them to random negative constants. Through empirical verification, we experienced better performance with randomizing the preferences and adopt this approach for most of our experiments.

Once the similarity matrix is constructed and provided as the primary input to the HAP algorithm, the network of nodes (data points) recursively transmits two kinds of intra-level messages between each node until a good set of exemplars is chosen. The first message is known as the "responsibility" message and the second as the "availability" message. The responsibility messages, ρ_{ij}^{l}, are sent at level l from data point i to data point j portraying how suitable node i thinks node j is to be its exemplar. Similarly, availability messages, α_{ij}^{l}, are sent at level l from data point j to i, indicating how available j is to be an exemplar for data point i. The responsibility and availability update equations are given in Eq. (1) and Eq. (2), respectively.

$$\rho_{ij}^{l>1} \leftarrow s_{ij}^{l} + min\,[\tau_i^l, -max_{ks.t.k\neq j}\{\alpha_{ik}^{l} + s_{ik}^{l}\}] \tag{1}$$

$$\alpha_{ij}^{l<L} \leftarrow min\left\{0, c_j^l + \phi_j^l + \rho_{jj}^l + \sum_{k\ s.t.k\notin\{i,j\}} \max\{0, \rho_{kj}^l\}\right\} \tag{2}$$

$$\alpha_{jj}^{l<L} \leftarrow c_j^l + \phi_j^l + \sum_{k\ s.t.k\neq j} \max\{0, \rho_{kj}^l\} \tag{3}$$

where L is the number of levels defined by the user and $l\in\{1,\cdots,L\}$. Eq. (3) is the self-availability equation which reflects the accumulated positive evidence that j can be an exemplar. The self-responsibility messages are updated the same way as the responsibility messages. To avoid numerical oscillation, the responsibility and availability messages are dampened by λ at every level l.

$$\rho_{ij}^{l} = \lambda\rho_{ij}^{l}(old) + (1-\lambda)\rho_{ij}^{l}(new) \quad (4)$$

$$\alpha_{ij}^{l} = \lambda\alpha_{ij}^{l}(old) + (1-\lambda)\alpha_{ij}^{l}(new) \quad (5)$$

HAP also introduces two inter-level messages that send messages between the different levels. These messages are denoted by τ in Eq. (6), which receives messages from the lower level and ϕ in Eq. (7), which receives messages from the upper level. At every level, the cluster preference c_i^l is updated using Eq. (8).

$$\tau_j^{l+1} = c_j^l + \rho_{jj}^l + \sum_{ks.t.k\backslash neq\ j} \max(0, \rho_{kj}^l) \quad (6)$$

$$\phi_i^{l-1} = max_k(\alpha_{ik}^l + s_{ik}^l) \quad (7)$$

$$c_i^l \leftarrow max_j(\alpha_{ij}^l + \rho_{ij}^l) \quad (8)$$

A variety of strategies can be employed to update the similarity matrix s_{ij}^l to vary level-wise. We have achieved good results by simply taking into consideration the cluster relationship of the previous level:

$$s_{ij}^{l+1} = s_{ij}^l + \kappa\, max_{j\ s.t.\ j\neq i}[\alpha_{ij}^l + \rho_{ij}^l] \quad (9)$$

where κ is a constant value within [0,1]. This updates the relation between data points in level l+1 by negatively increasing the similarity between points that belong to different clusters in level l and enforces the similarity between points that fall under the same cluster in level l.

After all messages have been sent and received, the cluster assignments are chosen, at every level, based on the maximum sum of the availability and responsibility messages as in Eq. (10). These cluster assignments can be used to extract the list of exemplars.

$$e_i^l \leftarrow argmax_j\{\alpha_{ij}^l + \rho_{ij}^l\} \quad (10)$$

These net message exchanges seek to maximize the cost of correctly labeling a point as an exemplar and gathering its representative members (a cluster). In Algorithm 1, we detail the pseudo-code implementation of HAP.

Algorithm 1 Hierarchical Affinity Propagation

Input: Similarity (S), Levels (L), Iterations, and λ
Initialize: $\alpha = 0,\ \rho = 0,\ \tau = \infty,\ \phi = 0,\ c = 0,\ e = 0$
for $iter = 1 \rightarrow$ Iterations **do**
for $l = 1 \rightarrow$ Levels **do**
Update ρ_{ij}^l (eq. 1) & Dampen ρ^l (eq. 4)
Update α_{ij}^l (eq. 2 & 3) & Dampen α^l (eq. 5)
Update τ_j^l, ϕ_j^l & c_j^l (eq. 6, 7 & 8)
Optional Update s_{ij}^l (eq. 9)
end for
end for
for $l = 1 \rightarrow$ Levels **do**
Update e_i^l (eq. 10)
end for

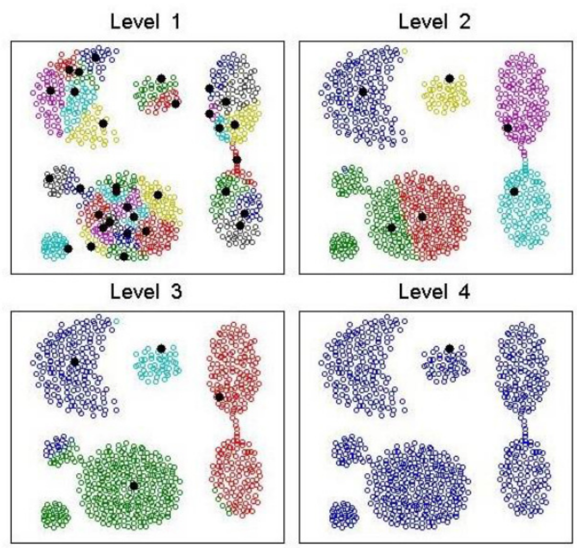


Fig. 3. Hierarchical Clustering with HAP

Figure 3 demonstrates the effectiveness of HAP to perform hierarchical clustering without predefining the number of clusters. The images are shown from the lowest level, top left image, to highest cluster level, lower right image. Once the clusters are formed on data-streams consisting of transportation-safety twitter messages, more detailed analyses can be performed to derive actionable intelligence based on particular safety emphasis areas. As example, the general public perception of safety-related issues can be extracted using sentiment analysis.

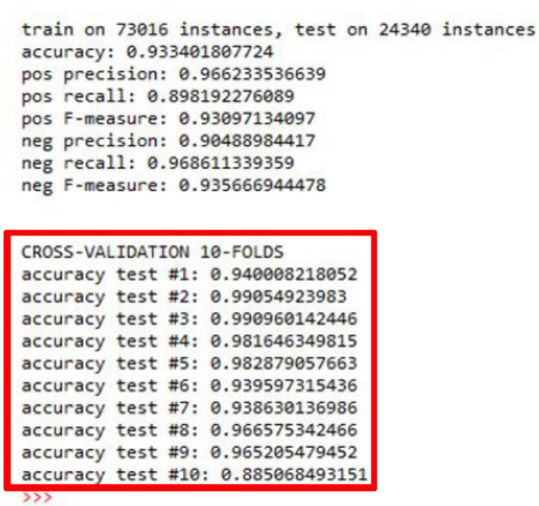
IV. SENTIMENT ANALYSIS COMPONENT

Once particular clusters are created using HAP, evaluation of human-perceived level of safety is carried out using sentiment analysis on each cluster of interest. Sentiment Analysis (SA) is a form of text classification for predicting attitudes or dispositions towards objects or persons. The task of SA can be briefly summarized as creating the prediction function γ(d)=c , where the classifier (γ) receives an input document (d) and produces or predicts a class (c), where c is the most likely representation of sentiment (e.g., positive or negative) conveyed by d . In our system, this process is streamlined by employing a bag-of-words model, where text documents are transformed into word (frequency) vectors that maintain each word in the document as a feature, but ignores other information such as special formatting or the order of the words (i.e., position) in the document. This model allows for a simplified approach for representing text documents as feature vectors that can be employed during training of the classification algorithms and for representing future documents (for prediction purposes) the same way.

Our initial effort employs a Naïve Bayes classifier to classify twitter messages that have been clustered using HAP to determine human-perceived quality of services as sentiment. The Naïve Bayes algorithm is presented in (11),

$$C \leftarrow argmax_{c_k \in K}\, P(C = c_k) \prod_i P(x_i|C = c_k) \qquad (11)$$

where C is the class (i.e., positive or negative) assigned to a given twitter message given a set of features x_i (i.e., each word in the message). Our initial dataset included a total of 97,356 labeled instances captured and filtered through Twitter's API 1.1. Our training and testing phase was performed using 10-fold cross-validation, where the dataset is divided into k=10 groups of samples, called *folds*. The prediction function was learned using k-1 folds, and the fold left out was used for testing. The results of the performance evaluation metrics are illustrated on Figures 4 and 5.



```
train on 73016 instances, test on 24340 instances
accuracy: 0.933401807724
pos precision: 0.966233536639
pos recall: 0.898192276089
pos F-measure: 0.93097134097
neg precision: 0.90488984417
neg recall: 0.968611339359
neg F-measure: 0.935666944478

CROSS-VALIDATION 10-FOLDS
accuracy test #1: 0.940008218052
accuracy test #2: 0.99054923983
accuracy test #3: 0.990960142446
accuracy test #4: 0.981646349815
accuracy test #5: 0.982879057663
accuracy test #6: 0.939597315436
accuracy test #7: 0.938630136986
accuracy test #8: 0.966575342466
accuracy test #9: 0.965205479452
accuracy test #10: 0.885068493151
>>>
```

Fig. 4. Output of n=10 cross-validation iterations

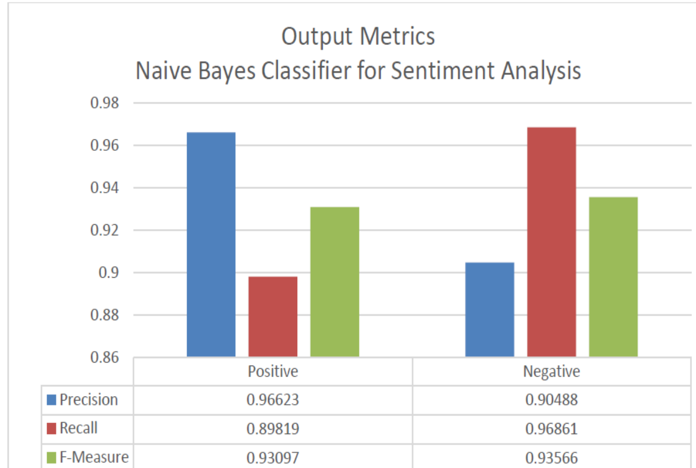


Fig. 5. Output of a Naïve Bayes classifier used for Sentiment Prediction of Tweets

From the performance metrics, the following is realized:

- Nearly every tweet that is *pos* is correctly identified as such, with 89% recall. This means very few ***false negatives*** in the *pos* class.
- A tweet given a *pos* classification is 96.6% likely to be correct. This leads to around 3% false positives for the *pos* label.
- Tweets that are *neg* are correctly classified. High recall throws a 4% false negatives for the *neg* label.
- Any tweet that is identified as *neg* is 90.8% likely to be correct. This means very few ***false positives*** for the *neg* class.

V. CONCLUSION AND FUTURE WORK

This paper presented work-in-progress for a big data analytics system for transportation safety management. Specifically, it presented the initial stages for determining human-perceived level of safety that can be extracted from social media data. Since the presented work deals with the general problem of determining human-perceived level of safety, there are several avenues for improvement and future work. Data cleansing and pre-processing poses a labor-intensive problem that needs to be resolved. Despite having used a focused twitter-search approach (related strictly to the area of transportation), many noisy messages were captured, expressing information unrelated to the desired topic. Many of the transportation-related messages were noisy and required pre-processing, which is a labor intensive task. In addition, other natural language scenarios in sentiment analysis need to be further investigated; the presence of sarcastic messages remains a difficult problem to tackle when using a bag-of-words approach. That is, sarcasm detection should be based on relation or dependence of two or more words, which clearly cannot be done under a bag-of-words model. This leads to some tweets containing sarcastic narratives that may correspond to a positive feeling, define a negative situation.

Future work includes extending the investigation to the following tasks:

- Delving on the contextual analysis of the messages in order to perform techniques of noise elimination.
- Data fusion and stream mining of real-time tweeter message flows, Waze, and current sensors used in transportation systems today.

REFERENCES

[1] National Safety Council, "NSC Motor Vehicle Fatality Estimates," Itasca, 2013.

[2] M. Peden, R. Scurfield, D. Sleet, M. Dinesh, H. A. Adnan, E. Jarawan and C. Mathers, "World Report on Road Traffic Injury Prevention," Geneva, 2004.

[3] P. Kruchten, "Architectural Blueprints - The "4+1" View Model of Software Architecture," IEEE Software, vol. 12, no. 6, pp. 42-50, 1995.

[4] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," Science, vol. 315, p. 2007, 2007.

[5] J. Hartigan and M. Wong, "Algorithm AS 136: A K-means clustering algorithm," Applied Statistics, pp. 100-108, 1979.

[6] L. Kauffman and P. Rousseeuw, Clustering by Means of Medroids, Fac. Univ, 1987.

[7] G. McLachlan and D. Peel, Finite Mixture Model, John Wiley & Sons, Inc, 2000.

[8] I. E. Givoni, C. Chung and B. J. Frey, "Hierarchical Affinity Propagation," CoRR, vol. abs/1202.3722, 2012.

[9]R. Guan, X. Shi, M. Marchese, C. Yang and Y. Liang, "Text Clustering with Seeds Affinity Propagation," IEEE Trans. Knowl. Data Eng., vol. 23, pp. 627-637, 2011.

[10] Scikit Developers, "scikit-lear.org," 2013. [Online]. Available: http://scikit-learn.org/stable/modules/cross_validation.html. [Accessed 29 May 2014].

[11] P.-N. Tan, M. Steinbach and V. Kumar, Introduction to Dana Minin, Boston: Pearson Education, Inc, 2006.

Building Blocks: Web Standards as Standard Resources in Robotics

Sekou L. Remy
School of Computing, Clemson University, Clemson, South Carolina, USA

Abstract—*Modern robotics often involves the use of web technologies as a means to cope with the complexity of design and operation. Many of these technologies have been formalized into standards, which are often avoided by those in robotics and controls because of a sometimes-warranted fear that "the web" is too slow, or too uncertain for meaningful control applications.*

In this work we argue that while web technologies may not be applicable for all control, they should not be dismissed outright because they can provide critical help with system integration. Web technologies have also advanced significantly over the past decade. We present the details of an application of a web server to perform open- and close-loop control (between 3Hz and 1kHz) over a variety of different network topologies. In our study we also consider the impact of a web browser to implement the control of the plant. Our results confirm that meaningful control can be performed using web technologies, and also highlight design choices that can limit their applicability.

Keywords: web standards, browser standards, robotics

1. Introduction

Modern robots are being applied in a dizzying array of tasks to solve a number of complex challenges. Unfortunately, as the complexity of the tasks increases, and as the complexity of the domain of available tools also increases, the development/programming of the robots may be just as difficult. The resulting complexity in robotic development requires a plurality of functional expertise (e.g. control systems, system administration, software engineering, mechanical engineering, etc.). As such, a significant obstacle in robotic development are the system integration challenges associated with getting tools and quickly getting these tools to work together.

Service-Oriented Computing [1], [2] is a paradigm that leverages the creation of basic blocks to quickly develop reusable applications [3]. The service-oriented approach is interesting in robotics because it readily provides an approach that permits subject matter experts to encapsulate their knowledge into useful building blocks that can then easily be used by others. ROS [4] and other service-oriented robotic offerings like those published in [5] are making effective strides, but more is needed.

We believe one path to accelerate service-oriented development in robotics builds upon advances in Networked Control Systems. These are control systems where networks provide the connections between spatially distributed system components [6]. In this work we discuss how the web, and standards developed for it that permit it to leverage the network, can more broadly be applied in robotics.

We begin with a brief characterization of the spectrum of web standards that can be applied in robotics. We then select one configuration of components selected from this spectrum and exploit it to demonstrate both the limitations and the opportunities for robotics and control.

Although not formally recognized in the literature, we believe that one source of objections relating to incorporating "the web" into robotics is related to timing variability and delay. These properties have a significant impact on feedback control systems. Therefore, exposing the impact that web standards can have on these properties is a valuable research contribution.

2. Background

Combining robotics and web technologies is not a new concept. As early as the mid-1990's robotic arms and other devices were being shared on the web for public and private use [7], [8]. At that time, sensing and actuation were severely constrained by the data rates of the existing technology. As such, teleoperation was, at best, tolerable; and meaningful automatic control tasks were not realistic applications of the technology.

Since that time, data rates on the Internet have increased by three orders of magnitude [9]. Accompanying such developments in Internet and computing technologies also have been advances in sensing and actuation. These advances have been built upon processing and storage, which are both faster and consume less power, and all this change provides opportunities for personal and industrial applications of web technologies – if they can be effectively harnessed in a principled manner. These opportunities can be advanced through the use of web standards.

2.1 Ecosystem of relevant standards

The Internet Engineering Task Force (IETF) and the World Wide Web Consortium (W3C) are two of the more visible bodies that are responsible for web (and Internet) standards. In fact, due to the omnipresence of the web in modern life, bodies like these increasingly function as standardization bodies for broader computing. For example, neither PNG (RFC2083) nor JPEG (RFC2045/2046) are data types that are only considered in web applications. They are common formats chosen for images in wherever computers are used. As another example, Bluetooth and Wi-Fi are two networking standards that are defined in the IEEE-802 suite of definitions. These standards are widely used in devices from televisions and printers, to the recently announced MYO, a Bluetooth connected EMG sensor.

In addition to data types and networking, there are protocols like HTTP (RFC 2616) that enable data communication over the multiple networks. Standardization has also enabled resources to be identified in a principled manner in this highly complex and dynamic environment. Such Uniform Resource Identifiers (URIs), defined in RFC3986, also include key=value pairs which readily enable data to be encoded for robotics and control use.

Finally, because these standards do not operate in a vacuum, the ecosystem around them is strongly influenced by their existence. For example, web servers and web browsers support many of these formal standards and although they are not formally sanctioned by any agency, they provide an informal standard that can be leveraged for our needs.

2.2 Approaches to leverage standards

We have presented the plurality of standards that are provided for the web, and in this section we describe how these standards can be combined in one configuration relevant to robotics, and subsequently we assess/quantify the characteristics of their use.

Specifically we leverage a web server, a RESTful web-service and clients written in python and javascript to show how standards that these components are based upon can be leveraged for autonomous control. Leveraging standards in this way provides a path to avoid the development of "yet another framework" that only runs on a given operating system, or with a given programming language, or only with a particular type of task. In this approach we demonstrate how operating frequency and networking conditions affect control. Our aim is to convince the reader that 1) there are many practical web standards that can be leveraged in robotics, 2) while "the web" may in general be the wild west of computing, there are configurations that can be effectively applied in relevant control.

3. Experiments

3.1 Configurations

In this work, we implement two types of clients, one written as a client-side javascript application and the other written in python. For both types of clients blocking requests are made to the server. This means that the client code halts until the server provides a response, or a time-out is triggered.

In the first portion of this work, the server's response is solely an acknowledgement that it has received data. This is a precursor to feedback control and is used to characterize the timing. Subsequently the server will respond with the updated state of the system coupled with it.

Table 1: Relevant properties for computers used in testing.

	OS	Firefox	Python	Processor
A	OSX 10.6.8	6.02	2.6.6	2.4GHz Core 2 Duo
B	Ubuntu12.04	3.6.12	2.6.5	2.4GHz Core i7
C	Ubuntu12.04	3.6.13	2.6.5	1.66GHz AtomN450
D	Ubuntu12.04	N/A	2.7.1	2.4GHz QEMU

Usability studies [10], [11] have long suggested human computer users demand response times on the order of 100-âĹŠ250ms. While typical web browsing does not require constant refreshing of content, this response time suggests that the underlying Internet infrastructure can permit information to change at rate of 4 Hz. As such we will run each client at 3.3 Hz, 10 Hz, and 100 Hz. The 100 Hz number is selected to show the potential of these resources in significant control applications, like quadrotor attitude control [12] or control of a PUMA robotic arm [13].

3.1.1 Locations and Computers

The experiments in this work were performed with clients running on four networked computers. The first machine, *Machine A*, is also running the web server. This machine is the constant during each of the experiments while the other clients vary slightly. *Machine A* is wirelessly connected to the network using 802.11g. The second machine, *Machine B*, sits on a wired gigabit ethernet connection to the network. The third machine, *Machine C*, also uses a wireless 802.11g network connection, but unlike *Machine A*, it is connected to a WPA2 encrypted network. WPA2 (also called WPA-802.1X) is a networking standard that defines security for wireless networks and is associated with increased network overhead. The fourth machine, *diego*, is housed in a remote data center. This machine was accessed in two ways, resulting in different network topologies. The first involved a wired gigabit ethernet connection, while the second leveraged a Virtual Private Network.

All communication in this work leverages HTTP. As indicated by the ping statistics (Table 2), blocking communication between client and server can take up to seconds to occur, but characteristically occurs on the order of tens of milliseconds. Packet loss is a possibility in network communication; however, the TCP provides a reliable connection between nodes and will retransmit lost packages (invariably adding delay).

A five second timeout is used in the python clients, and the javascript clients in Firefox are also documented to by default also have a five second communication timeout that is known as an "unresponsive script" timeout.

Table 2: Statistical summary of the response packets received from ping commands between the server (*Machine A*) and each of the machines used to run the clients. diego[2] includes an off-campus, VPN facilitated communication.

	min(ms)	**avg**(ms)	**max**(ms)	σ(ms)	**loss**
A	0.031	0.108	0.625	0.069	0.0
B	1.360	3.376	4.539	0.484	0.0
C	4.896	104.188	3033.374	336.043	1.6%
D[1]	0.943	3.155	13.42	0.752	0.9%
D[2]	39.01	53.14	145.2	21.75	0.0

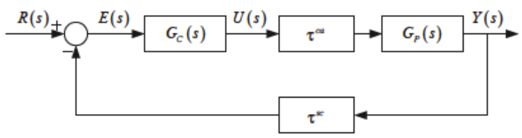


Fig. 1: Closed-loop control system example.

3.1.2 Data Collected

In the message sent from client to server in each of the runs, we log the time that the message was sent and the time it was received. We do not perform clock synchronization between the two machines, so all reported results are based on relative time changes with the candidate reference points as the first noted time stamp for each of the time sources.

3.2 Closed-loop Control

To explore the impact of timing on closed-loop control, the client and the server are coupled with digital implementation of a motor controller and a D.C. motor. The transfer function for the controller and the plant are selected from [14]. In the original work these authors investigated the performance degradation introduced by delays in the control loop, and they are used for the same reason here.

The transfer function implemented in the client is defined as follows:

$$G_C(s) = \frac{a_0 s + a_1}{s}, \tag{1}$$

where $a_0 = 0.1701$ and $a_1 = 0.378$. In the server, the plant's transfer function is:

$$G_P(s) = \frac{b_0}{(s + c_0)(s + c_1)}, \tag{2}$$

and $b_0 = 2029.826$, $c_0 = 26.29$, and $c_1 = 2.296$.

This closed-loop control system, captured in Fig. 1, performs PI control on the plant, and is slightly different from the cited source. In this original work, τ^{ca} and τ^{sc} (the delay between controller and actuator, and between sensor and controller) are experimental parameters. In this work, however, these two delay values are influenced only by the location of the client. Another difference is that the values of τ^{ca} and τ^{sc} are only weakly correlated. While data often travels the same path to and from the server, it is not guaranteed to do so. Furthermore, even if the path remained consistent, the conditions along the way can vary from time to time. For this reason, the assumption that $\tau^{ca} = \tau^{sc}$ does not hold.

Table 3: Percentage of packets sent from javascript clients that arrived at the server (*Machine A*) at the same time as the previous packet in the sequence during 600 seconds of client operation.

	100Hz	**10Hz**	**3.3Hz**
A	0.2494	0.0	0.0
B	0.0746	0.0	0.0010
C	0.0882	0.0051	0.0005

For both client and server, to discretize the transfer function, the backward Euler method is applied. This method, widely used in commercial control applications, is one integral based, numerical approximation method. It is applied by using the substitution formula in (3),

$$s = \frac{z - 1}{zh}, \tag{3}$$

where h is the simulation timestep. In this work, the same timestep is applied in both the controller and the plant. Values of $h = 3ms$ and $h = 1ms$ were both applied in this work.

4. Results

While TCP guarantees in-order and lossless communication between end points, there is no guarantee that the web server will process all messages in order. Due to factors such as packet buffering, queuing, message retransmission, and operating system scheduling, it is possible for sequential network packets to be provided to the server at the same time. When messages arrive at the server at the same time, the server spawns threads to process them. No guarantees can be made on which threads complete their processing first however. As such, it is possible for messages that arrive at the same time to be applied out of order in spite of TCP's sequential guarantees.

As indicated in Table 3, this occurrence is affected by both the operating frequency of the client and the network connection to the server. The effect of occurrences like these is that control commands can be issued by the server out of order. This means that control noise is introduced in the data generated by the client.

While this effect is magnified by higher operating frequencies, its impact is also marginalized with faster control. Since higher operating frequencies result in shorter time between control actions, there is also less time for the system dynamics to develop. In this work, this effect is noticed as additional noise in the inter-departure times and the existence of zero values in the inter-arrival times. Inter-arrival times should not be less than the inter-departure times, which in the ideal case are the inverse of the operating frequency.

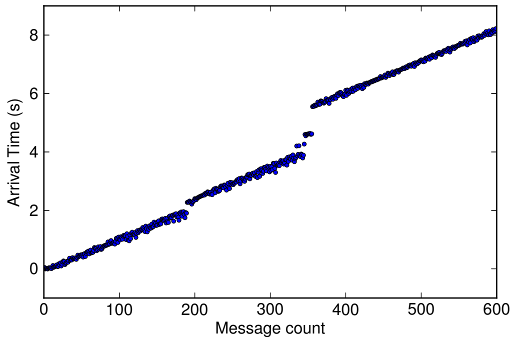
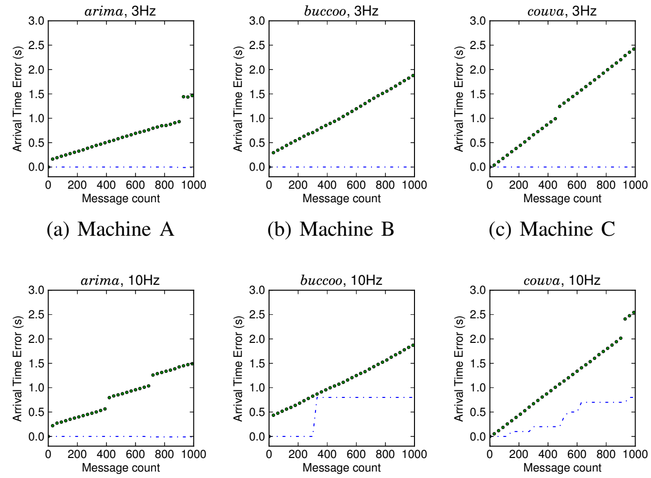


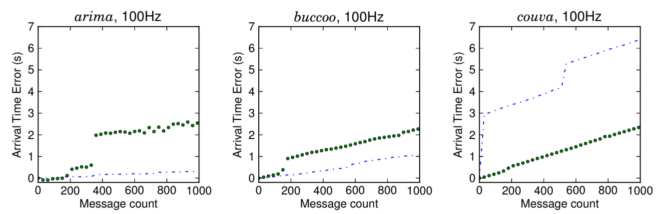
Fig. 2: First 600 arrival times for messages from the javascript client running at 100 Hz on *Machine A*.



(a) Machine A (b) Machine B (c) Machine C

(d) Machine A (e) Machine B (f) Machine C

Fig. 3: The difference in the time stamps of generated messages from python (blue) and javascript (green) operating at 3.3Hz on *Machine A* (3a), on *Machine B* (3b), and on *Machine C* (3c). Figures 3d-3f are for 10Hz control on these machines as well.



(a) Machine A (b) Machine B (c) Machine C

Fig. 4: The difference in the time stamps of generated messages from python (blue) and javascript (green) operating at 100Hz clients on each of the three machines.

4.1 Timing with 100 Hz Javascript Client

In figure 2 we show the first 600 samples of the time series generated by the javascript client running on *Machine A*. This figure captures variability, delay and a dilation effect (the altered slope) that is likely to be related to processing overhead.

The last two of these terms, overhead and transmission delay, both have long-term impact on the number of messages that are transmitted during control. Specifically, they both result in increased system latency. The 600th message should have been sent six seconds after the first message. This graph shows that it was sent two seconds late. In the first six seconds, just shy of 400 messages were transmitted. Over the ten minutes of the client's operation, only 55 thousand messages were transmitted, better than the effective rate observed during the first six seconds, but still short of the 60 thousand messages that should have been sent.

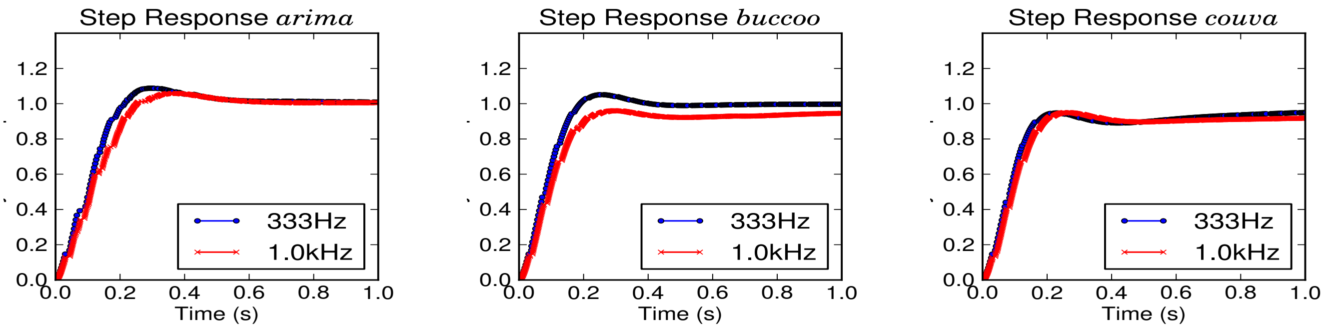
If open-loop control was performed under such conditions, this would mean that after 600 seconds of operation, the controller would be issuing control that is roughly 50 seconds delayed. For every ten seconds of control, the controller falls an additional second behind. For applications with long standing interaction, this would be quite problematic. When closed-loop control is considered, however, this is diminished. After ten minutes of operation, the controller would still be operating on the most recently received input, which according to the data from traceroute and ping commands (See Table 2), should only be delayed on the order of tens of milliseconds.

4.2 Impact of client location

To explore how the timing vagaries are affected by the selection of different clients, we next consider how the messages generated by the clients compare to the desired reference timeline. We test this on each of the three machines and first consider javascript and python clients running at 3.3Hz.

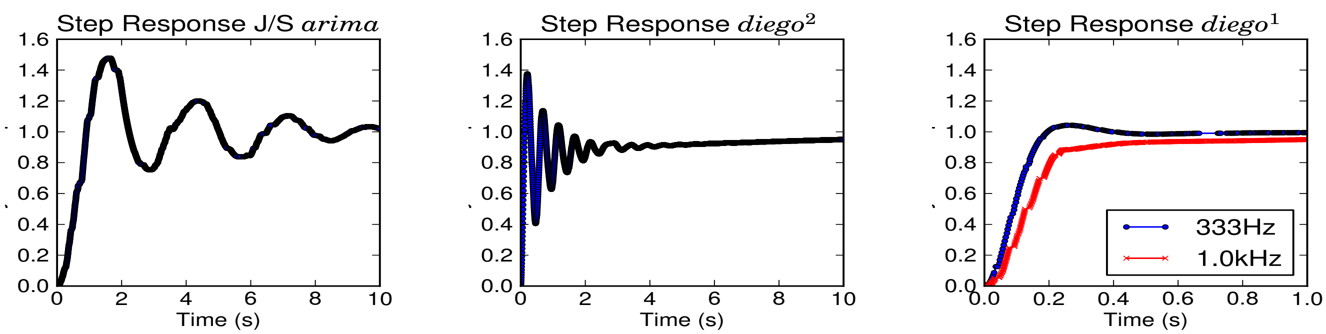
In figure 3, the plots of error for the respective clients are shown. In all of these cases, the python client effectively generates messages that are received on the desired timeline. This is evident because there is little difference between the timestamps for the client messages and the ideal values (zero slope). What should also be evident from these figures is that the javascript clients all diverge from the reference with time (non-zero slope).

The same basic trend is observed at higher operating frequencies (See figure 4). At these higher frequencies, the impacts of the network delay and network jitter are more pronounced because the messages are being transmitted more frequently. At the same time, the impact of noise in the client timing infrastructure is also more pronounced. One millisecond of error is more significant in ten millisecond intervals than it would be when that interval is 100 or 300 milliseconds. Messages should have the same base



(a) Machine A (b) Machine B (c) Machine C

Fig. 5: Unit step responses for controllers and plants running under different network locations.

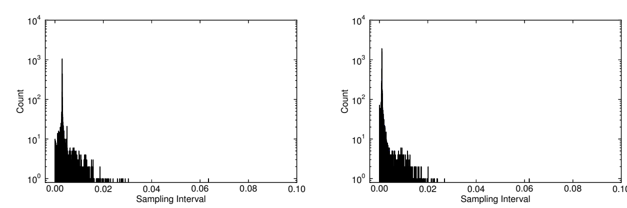


(a) Machine A (b) Machine D^2 (c) Machine D^1

Fig. 6: Unit step responses operating at 333Hz for javascript client on *Machine A* (6a), and for python clients on *diego* connected via a VPN from off-campus (6b) and over the campus LAN (6c).

transmission delay, but the effects of queuing and conflict resolution for the shared medium are now amplified. (Recall that the server's acknowledgement is also guaranteed to be sent back to the client.) These effects are evident in the non-zero differences from the reference.

At 10Hz, even though the python client (See Figures 3d-3f) has non-zero differences from the reference, the deviations are largely isolated to impulses (i.e. errors injected at instances) not the continuous deviation that appears constant for javascript clients. For these clients, there is no significant divergence as the slopes of data (in blue) are essentially zero. As the highest operating frequency, 100Hz, such stability no longer exists in the data (See Figures 4a-4c).



(a) 333Hz (b) 1kHz

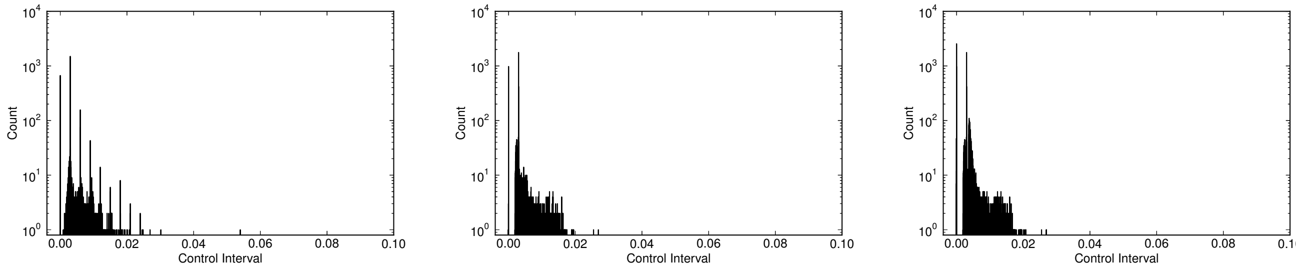
Fig. 7: Sample intervals observed for the server operating with 0.001s intervals and 0.003s intervals (Figure 7a and 7b respectively).

4.3 Closed-loop Control

From the figures presented thus far, it is clear that undesirable timing characteristics exist between the clients and the server. To contextualize the impact of these, closed-loop control was performed over the network and facilitated by the web server, URI encoded queries and a web browser to run javascript versions of the client. The operating frequencies in these closed-loop cases are higher than the 100Hz previously considered enabling comparison with published results.

As the graphs in Figure 5 confirm, even in the light of these timing challenges, closed-loop control with python controllers on *Machine A*, *Machine B*, and *Machine C* were all successfully demonstrated.

The performance is in keeping with [14] and even when the operating frequency is increased to 1 kHz, effective control was performed. Recall that this controller was not



(a) Machine A (b) Machine B (c) Machine C

Fig. 8: Distribution of control intervals observed for each controller operating at 333Hz (each should be 0.003s).

designed for network use, and makes no accommodation for the timing challenges of distributed control. When the javascript controller is used however, the settling time of the system increases ten-fold (see Figure 6a). The system remains stable, but would not be practical for use where response times on the order of fractions of seconds are required. It may be possible to leverage Kalman filters to avoid this issue, but it is not yet clear what circumstances should warrant such efforts.

Surprisingly, as seen in Figure 6b, when a python controller is run from an on-campus location, to an off-campus plant (server), the settling time was only quadrupled. If that same location is used and the plant is moved on-campus, its operation is in keeping with performance on *Machine A*, *Machine B*, and *Machine C* for both 333Hz and 1KHz control.

Plainly put, these results confirm that meaningful closed-loop control, in this case control of D.C. motor with time constant $30ms$, is only negatively impacted the use of web standards when a javascript client is used. The javascript client, which runs in-browser, is crippled because it shares the same thread for control and the other core functionality of the browser's operation. The usefulness of this resource for high-speed control is thus significantly hindered by this operating constraint.

To reveal why the system is able to successfully operate under these circumstances, we explore how the intervals in which new control signals are generated are affected by client/server locations.

As can be seen in Figure 7, the sampling intervals on the computers are quite noisy. The higher the frequency, the larger the variation in the intervals. As expected, the means of these intervals are $3.0ms$ and $1.004ms$ while their standard deviations are $2.984ms$ and $3.705ms$ respectively.

It is known that real-time operating systems have merit in control, but these plots highlight precisely how. Even before considering the variations introduced by the use of the network or web standards, there is noise injected into system control. The reduction in the number of negative and zero-length intervals also evident in this figure is associated with the additional overhead to calculate control values, a factor not previously present for the data in Table 3.

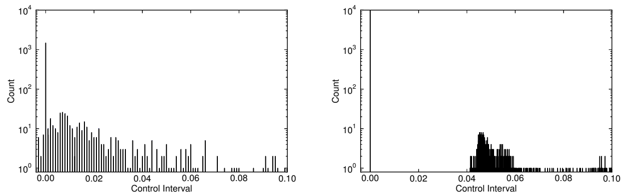
The impact of variability can also be seen when comparing the histograms in Figure 8. For the client that is closest to the server (Figure 8a), the impact of the controller sampling frequency is the dominant factor influencing the control intervals. In this case, since the sample time in on the order of the transmission time, it is possible, especially in the face of the sampling noise, to receive multiple control signals during a single sampling interval. It is also possible to send multiple control signals before the response is received from the plant (and new control values are generated). These two factors account for the strong presence of periodic intervals that are multiples of the sampling time, and also for zero-length intervals.

As the transmission time for control signals increased (Figures 8b and 8c), the likelihood of sending multiple control signals before receiving a response increases. At the same time, the possibility of the plant receiving multiple control signals in a sample interval decreases. These two changes are appropriately reflected in these histograms (flatter noise profile, and more zero-length intervals), as well as in Figure 9b.

These graphs all suggest that the use of a desktop operating system, and communication with very high transmission latency have a more significant impact than the use of the web server, and URI encoded control (and responses).

The timing challenges of using javascript at this frequency have already been stated. Neither the browser, nor the scripting language was designed to operate in this manner, and it is clear when one considers the wide distribution of control intervals associated with $3ms$ intervals (See 9a). Design notwithstanding, this histogram shows support for avoiding control with time intervals on the order fractions of seconds. This guideline should be reconsidered within 2 years, as modern browsers are transitioning to different interaction paradigms and implementations.

Considering the histogram for control intervals to an off-



(a) Machine A (b) Machine D^2

Fig. 9: Undesirable effects of client type and network delay on the desired control interval (0.003s). Figure 9a is was measured for the javascript client running on *Machine A* while 9b is for the python client running on *diego* connected via a VPN from off-campus.

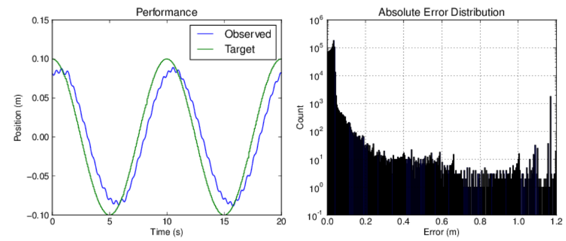


Fig. 10: Closed-loop control of an unstable system (5 hours 45 mins of data). The amplitude of the target is 0.1m, and the median absolute error is 0.03m.

campus controller (Fig. 9b), the impact of transmission delay at minimum on the order of $40ms$ is evident. As previously indicated, the increased transmission time results in more zero-length intervals and although not even the target sampling rate is evident, control can still be stably demonstrated in the system.

All of the presented data are for relatively short time intervals (on the order of minutes). This time scale facilitated the initial focus and confirmed that the infrastructure was reasonable for the task. This provides the needed validation to continue to quantify the impact of these standards in more challenging conditions. When using networks, especially wide area networks, or the Internet, the jitter and delay can vary radically over larger timescales. Figure 10 is generated from the control of the plant over a VPN connection between Oklahoma and South Carolina, using the public Internet. The data show that while the impact of noise is evident (upto 1m of error), most of the error is less than 30% of the target amplitude.

5. Conclusion

From these results we can conclude that relevant control can be performed at operating frequencies meaningful in robotics. Further study is needed with systems that are unstable to characterize the opportunities for other relevant classes of robotic systems. Preliminary work has begun with the ball on plate.

This work shows that by leveraging resources and standards long pervasive in networked control environments, that exciting opportunities to connect to high-level distributed control systems, with little additional investment by researchers and developers. Web standards and browser standards are powerful, yet underutilized tools in many computing applications.

To enable these resources to be more widely appreciated, we advocate that meaningful collaboration must occur across disciplines and many practical activities attempted to permit researchers to learn each other's languages and to recognize each other's needs. In this way our fields can become better equipped to address the many complex and interesting challenges that we have long dreamed to address.

References

[1] M. Huhns and M. Singh, "Service-oriented computing: key concepts and principles," *Internet Computing, IEEE*, vol. 9, no. 1, pp. 75 – 81, jan-feb 2005.

[2] M. Blake, S. Remy, Y. Wei, and A. Howard, "Robots on the web," *Robotics Automation Magazine, IEEE*, vol. 18, no. 2, pp. 33 –43, june 2011.

[3] M. P. Papazoglou, P. Traverso, S. Dustdar, and F. Leymann, "Service-oriented computing: a research roadmap," *International Journal of Cooperative Information Systems*, vol. 17, no. 02, p. 223, 2008.

[4] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: An Open-Source Robot Operating System," in *ICRA Workshop on Open Source Software*, 2009.

[5] Y. Chen, Z. Du, and M. Garcia-Acosta, "Robot as a service in cloud computing," in *Service Oriented System Engineering (SOSE), 2010 Fifth IEEE International Symposium on*, june 2010, pp. 151 –158.

[6] H. Lin, G. Zhai, and P. Antsaklis, "Robust stability and disturbance attenuation analysis of a class of networked control systems," in *42nd IEEE Conference on Decision and Control*, vol. 2, Dec. 2003, pp. 1182 – 1187.

[7] K. Y. Goldberg, M. Mascha, S. Gentner, N. Rothenberg, C. Sutter, and J. Wiegley, "Destop teleoperation via the world wide web," in *ICRA*, 1995, pp. 654–659.

[8] P. Backes, G. K. Tharp, and K. S. Tso, "The web interface for telescience (wits)," in *in Proc. IEEE Int. Conf. Robot. Automat*, 1997, pp. 411–417.

[9] J. Nielsen, "Nielsen's Law of Internet Bandwidth," *http://www.nngroup.com/articles/nielsens-law-of-internet-bandwidth/*, Apr. 1998, accessed 2013-01-02. [Online]. Available: http://www.nngroup.com/articles/nielsens-law-of-internet-bandwidth/

[10] R. B. Miller, "Response time in man-computer conversational transactions," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, ser. AFIPS '68 (Fall, part I). New York, NY, USA: ACM, 1968, pp. 267–277. [Online]. Available: http://doi.acm.org/10.1145/1476589.1476628

[11] J. Nielsen, *Usability Engineering*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.

[12] G. M. Hoffmann, S. L. Wasl, and C. J. Tomlin, "Quadrotor helicopter trajectory tracking control," in *In Proc. AIAA Guidance, Navigation, and Control Conf*, Aug. 2008.

[13] S. Hayati, T. Lee, K. Tso, P. Backes, and J. Lloyd, "A unified teleoperated-autonomous dual-arm robotic system," *Control Systems, IEEE*, vol. 11, no. 2, pp. 3–8, Feb. 1991.

[14] Y. Tipsuwan and M.-Y. Chow, "Control methodologies in networked control systems," *Control Engineering Practice*, vol. 11, no. 10, pp. 1099–1111, 2003. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0967066103000364

A Sensor-Oriented Information System Based on Hadoop Cluster

Yu Liang* and Chao Wu
Department of Computer Science and Engineering, University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA

Abstract - *In order to obtain a real-time situational awareness about the specific behavior of target-of-interests out of huge-scale sensory data-set, this proposed work presents a generic sensor-oriented information system based on Hadoop cluster (SOIS-Hadoop). NoSQL database is used to store and manage the heterogeneous sensory data; Hadoop/MapReduce programming paradigm is employed to optimize the parallelism of data retrieval and analysis. Using two representative sensor-oriented information processing and analysis problems as examples, the mechanism of data analytics based on modeling and simulation for sensory data is also presented in this paper.*

Keywords: *sensor-oriented analysis*, *Hadoop cluster*, *NoSQL*, *modeling and simulation*, *data analytics*, *mathematical model*.

1 Introduction

It is an extremely computationally intense, labor-intensive and highly unreliable job to obtain a panoramic, timely, trusted understanding about the observed behavior of target-of-interest (TOI) [1-3]and its future status by exploiting networked sensor assets (autonomous, heterogeneous and multi-layer sensor nodes). Featured with massive volume, high generation velocity, large variety of formats, and great value, the sensor-oriented information system can be regarded as a typical big-data problem, which is conventionally implemented using Hadoop platform[4].

This project proposes a generic sensor-oriented information system based on Hadoop cluster (SOIS-Hadoop) to monitor and analyze the specific behavior of target-of-interest (TOI) according to persistent surveillance sensory data. The addressed SOIS-Hadoop has the following features:

- Maintaining the consistency of difficult target-of-interest via trajectory prediction based on mathematical modeling and simulation[5-7].

- Using smart sensor network techniques such as collective control and self-optimization to achieve the optimal observations about target-of-interest.

- Employing Hadoop/MapReduce programming paradigm [4] to handle the processing and storage of huge-scale sensor data over cluster platform.

The remainder of the paper is organized as follows: Section 2 discusses the hardware infrastructure of the SOIS-Hadoop system; Section 3 discusses the software framework of the SOIS-Hadoop system; Section 4 briefly introduces the flowchart of the system; Section 5 uses two representative examples to demonstrate the data analytics according to sensory data extracted from Hadoop; Section 6 summarizes the effort.

2 Hardware Architecture of Hadoop Xen Clusters

As an efficient and timely strategy for processing large-scale data, Hadoop allows developers to create distributed applications running on clusters of computers. This infrastructure can then be leveraged to tackle very large data sets--by breaking up the data into "chunks" and coordinating the processing of the data out into the distributed, clustered, environment.

Fig. 1 shows the hardware-configuration of Hadoop based sensory data processing and analysis system. The proposed system consists of three hardware components:

- Data acquisition and pre-processing based on mobile computing platform (e.g., iPhone, laptop, etc.). Sensory data may be acquired by multiple sensors at the same time. Pre-processing indicates translating stream sensory data such as video data into semi-structured format data such as XML[5-8].

- Data storage and analysis based on Hadoop cluster, which is formulated using multiple inter-connected Xen virtual stations.

- Data analysis and visualization. The component provides a user's interface.

The proposed hardware system is completely constructed over internet, the bandwidth and reliability will determine the accuracy of the analysis outcome.

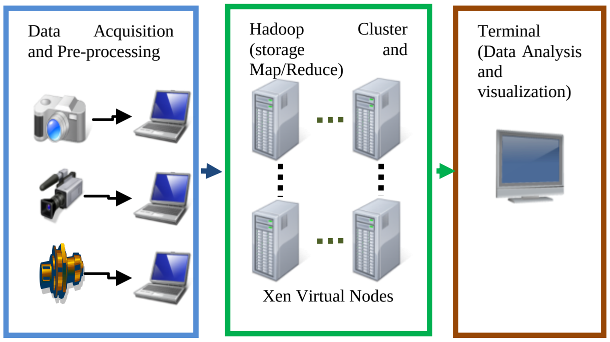


Fig. 1 Hardware Configuration of SOIS-Hadoop (dash-line indicates those internet link).

3 Software Infrastructure Of SOIS-Hadoop

In order to efficiently detect, track, understand and evaluate the behavior of target-of-interests (particularly for those difficult target) activity according to persistent surveillance sensory data, this project investigates a generic sensor-oriented information system (SOIS) based on Hadoop/MapReduce (SOIS-Hadoop) programming paradigm [4] is given in Fig. 2. As illustrated in Fig. 2, geographic information system and persistent surveillance sensory data constitute two major important inputs for SOIS-Hadoop; Hadoop/MapReduce module processes, integrates, analyzes and stores the sensory data over a distributed cluster platform in parallel; analytics about sensory data is obtained based on the mathematical modeling about the observed behavior of target-of-interest.

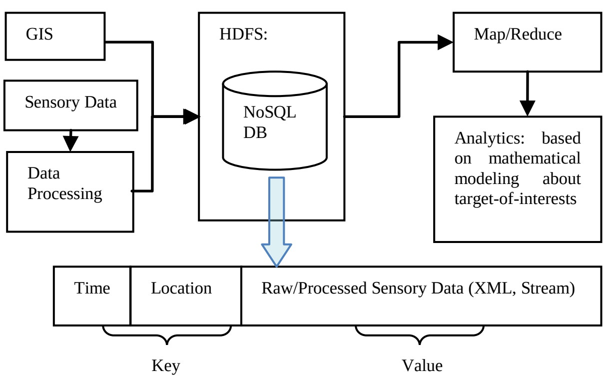


Fig. 2: Infrastructure for SOIS-Hadoop.

Sensor-oriented information analysis is generally featured with huge amount of data processing, query-type work load, Hadoop/MapReduce is employed in this work because it provides (1) distributed storage model for huge-volume sensory data, and (2) large-scale data parallel processing framework [4]. NoSQL database [9] is used to store the sensory data, most of which are semi-structured or none-structured. Compared to traditional database format, NoSQL database are applicable for dynamic data structure and has superior scalability because it does not need to category and parse the sensory data into fixed format.

4 A Generic Flowchart For Sensor-oriented Information Analysis System

Fig. 3 shows a generic flow-chart about the sensor-oriented information analysis system [9]. Among the modules described in Fig. 2, geographic information module defines the geometry configuration of the scene; mathematical model about the expected behavior of TOI is generally given in the form of partial differential equations, which is used to anticipate the evolution of observed behavior about TOI; persistent surveillance sensory data is directly acquired from sensors.

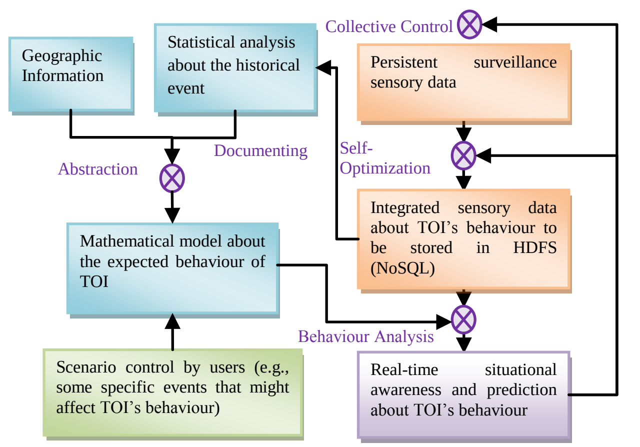


Fig. 3: A generic flow-chart about sensor-oriented information analysis system.

Fig. 3 also illustrates that the implementation of sensor-oriented information analysis system consists of following two threads: (1) Formulating the mathematical model (e.g., spatial- and temporal-dependent partial differential equations) using historical sensory data about the expected behavior of target-of-interest (TOI) [1, 2, 5-8, 10]. (2) Processing and integration of observed sensory data. A situational awareness is derived from these two threads [5-8]. Then the situational awareness will reversely guide the self-optimization of data-collection and cooperative control of sensor nodes so as to generate more accurate understanding about external events.

The mathematical model about the expected behavior of target-of-interest is formulated according to the geographic information and the historical records about target-of-interests. The core calculations corresponding to mathematics modeling include (1) statistical analysis about the historical sensory data stored in HDFS, and (2) numerical solution to the mathematical simulation (e.g., using finite element method [10, 11] to solve the temporal and spatial-dependent partial differential equations).

Processing of persistent surveillance video data includes the following operations: (1) acquisition of video data; (2) segmentation, which extracts pixels of target-of-interests (TOI) from background; (3) isolation of TOIs out of noise or other moving objects; (4) translation of optical behavior features (i.e., the velocity and position of moving targets within the sensor coordinate system) of detected pedestrians into their actual geographical features (i.e., the velocity and position of moving targets within the geographic coordinate system); (5) documentation, which posts the output in a format suitable for post-processing and includes position, velocity, and track. The implementation of both threads is highly computation and storage-intensive.

5 Data Analytics Based On Mathematical Model

In the context of sensor-oriented analysis, data analytics of sensory data aims to disclose specific behavior of target-of-interest (TOI) out of sensory data [5-8]. For example, it is a significant task to detect those speeding or wrong-way vehicle out of surveillance video on the road; therefore a description about the expected (or normal) traffic flow is needed so that those abnormal vehicles can be identified. The expected behavior (i.e., normal behavior) of TOI is commonly modeled using micro-scale or macro-scale method. Micro-scale method, which is also called agent-based method, provides a detailed formulation about the behavior of TOI while suffers from inhibitive computational cost and accumulated numerical error. Macro-scale method generally uses time- and space-dependent partial differential equations (PDE) to formulate the expected behavior of TOI. In this paper, micro-scale method is used to simulate the aggregation of carp and macro-scale method is used to identify anomalous vehicle on the road.

5.1 Aggregation of Carp

It is known that Asian carps are causing serious damage to the area' fresh-water ecosystem because they out-compete native fish for food and habitat, lower water quality by killing off sensitive organisms like native freshwater mussels. In order to control the populations of Asian carps, a molecular-dynamics based mathematics model (a micro-scale method) is presented to formulate the aggregation of Asian carps in this work.

The force-field among carps, which is derived from empirical sensory data, is formulated according to the pairwise interaction between neighboring carps. The pairwise interaction U_{ij} is defined using modified van der Waals forces, where the corresponding potential function U_{ij} between carp-i and carp-j is defined by the Equation (1).

$$U_{ij} = \begin{cases} 4\varepsilon\left[\left(\frac{\sigma}{\|r_{ij}\|}\right)^{12} - \left(\frac{\sigma}{\|r_{ij}\|}\right)^{6}\right] & \left(\|r_{ij}\| < R_s\right) \\ 4\varepsilon\left[\left(\frac{\sigma}{R_s}\right)^{12} - \left(\frac{\sigma}{R_s}\right)^{6}\right] & \left(R_s \le \|r_{ij}\| \le R_h\right) \\ 4\varepsilon\left[\left(\frac{\sigma}{\|r_{ij}\|-R_h+R_s}\right)^{12} - \left(\frac{\sigma}{\|r_{ij}\|-R_h+R_s}\right)^{6}\right] & \left(R_h < \|r_{ij}\|\right) \end{cases} \quad (1)$$

where:

$$r_{ij} = X_i - X_j \quad (2)$$

and

$$\sigma = \frac{R_s}{\sqrt[6]{2}} \quad (3)$$

In Equation (1), R_s, R_h and R_k are illustrated in Fig. 4. $\|r_{ij}\|$ indicates the distance between two neighboring carps. σ and ε are constant coefficient for van der Waals forces. It should be remarked that the moving orientation, water flow velocity and blind zone is not considered in the formulation of Equation (1).

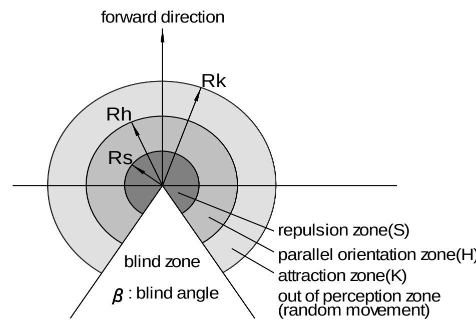


Fig. 4: Interaction zones between neighboring carps

As illustrated in Fig. 5, given the initial status of carp school (Fig. 5a), the future status of carp school can be derived from the sensory data using molecular-dynamics mathematics model (Fig. 5b). Based on the preliminary simulation results, the motivations of fish aggregation, such as foraging advantages, reproductive advantages, predator avoidance, or hydrodynamic efficiency, can be disclosed.

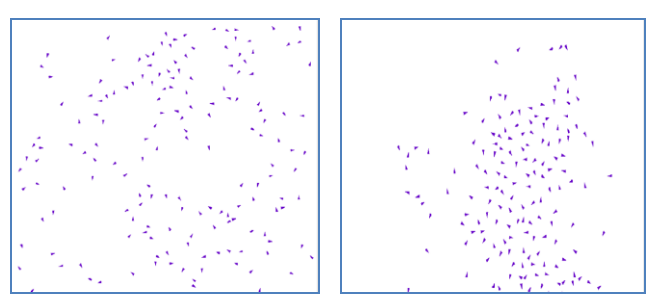


Fig. 5: Aggregation of carp: (a) initial status of carp school acquired via sensors; (b) aggregation predicted by data analytics.

5.2 Vehicle Traffic Analysis

In vehicle traffic analysis, the expected traffic flow is formulated using temporal and spatial-dependent partial differential equations [5]. The governing equations for traffic flow are defined by the following partial differential equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho V) = 0 \tag{4}$$

$$\rho(X,t) = \frac{r}{V_{\max} A(X) \log A(X)} \tag{5}$$

$$\rho \frac{DV}{Dt} = -\nabla P + \mu \nabla^2 V + \rho g \tag{6}$$

Where $\rho(X,t)$ is the number of vehicles over unit length, $V(X,t)$ is the expected velocity of the vehicle, $V_{\max}$ is the speed limit, $A(X)$ is the cross-section width (or bandwidth) of the road. Equation (4) is derived from conservation of mass. Equation (5) ensures that traffic flow slows down up at nozzle and keeps constant speed at fork.

Fig. 6a shows the pressure field and velocity field resulted from the solution of governing equations. The boundary conditions and the coefficient equations for governing equations are obtained according to historical traffic data. Using the density field, velocity field and pressure field about traffic flow as reference, the observed vehicles can be measured and evaluated. As illustrated in Fig. 6b, a blue-circled car is identified as "abnormal" because its movement is inconsistent with expected traffic is identified and noted using red-line.

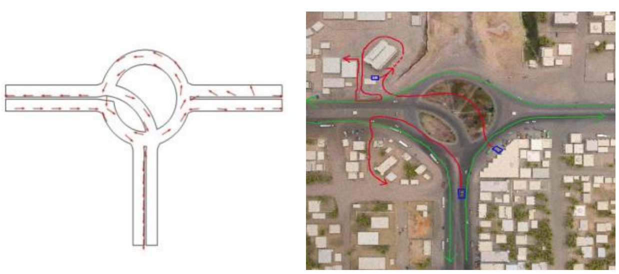


Fig. 6 (a) expected traffic flow; (b) identified anomalous vehicle (circled in blue color) with reference to expected traffic flow.

6 Summary

A simple SOIS-Hadoop platform has been set-up and applied in a variety of practical problems such as the spread of epidemics diseases [7], vehicle traffic analysis[2, 5, 8], anomalous pedestrian detection [6], etc. Some preliminary while promising outcomes has achieved.

Our future work will focus on the high-performance storage and retrieval of sensory data in NoSQL database.

7 Acknowledgment

This work is jointly sponsored by the National Science Foundation (NSF) with proposal number 1240734 ("*A Design Proposal for the Center of Cyber Sensor Networks for Human and Environmental Applications*") and 1111542 ("*RI: Large: Collaborative Research: A Robotic Network for Locating and Removing Invasive Carp from Inland Lakes*").

8 References

[1] S. Fernandes, Y. Liang, S. Sritharan, X. Wei, and R. Kandiah. "Real Time Detection of Improvised Explosive Devices using Hyperspectral Image Analysis"; in *2010 IEEE National Aerospace and Electronics Conference (NAECON 2010)*.

[2] S. Fernandes and Y. Liang. "Chipping and segmentation of target of interest from low-resolution electro-optical data"; in *SPIE Defense, Security, and Sensing*, pp. 87440R-87440R-8, 2013.

[3] Y. Liang, W. Melvin, S. Sritharan, S. Fernandes, and B. Barker. "A Crowd Motion Analysis Framework Based on Analog Heat-Transfer Model"; *American Journal of Science and Engineering (In-Press)*, 2013.

[4] R. C. Taylor. "An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics"; *BMC bioinformatics*, vol. 11, p. S1, 2010.

[5] Y. Liang, M. Henderson, S. Fernandes, and J. Sanderson. "Vehicle tracking and analysis within a city"; in *SPIE Defense, Security, and Sensing*, pp. 87510F-87510F-15, 2013.

[6] Y. Liang, W. Melvin, S. I. Sritharan, S. Fernandes, and D. Barker. "CMA-HT: a crowd motion analysis framework based on heat-transfer analog model"; in *SPIE Defense, Security, and Sensing*, pp. 84020J-84020J-13, 2012.

[7] Y. Liang, Z. Shi, S. I. Sritharan, and H. Wan. "Simulation of the Spread of Epidemic Disease Using Persistent Surveillance Data"; *Simulation*, p. 1, 2010.

[8] J. Sanderson and Y. Liang. "No-reference image quality measurement for low-resolution images"; in *SPIE Defense, Security, and Sensing*, pp. 874404-874404-17, 2013.

[9] K. Grolinger, W. A. Higashino, A. Tiwari, and M. A. Capretz. "Data management in cloud environments: NoSQL and NewSQL data stores"; *Journal of Cloud Computing: Advances, Systems and Applications*, vol. 2, 2013.

[10]Y. Liang, J. Weston, and M. Szularz. "Generalized least-squares polynomial preconditioners for symmetric indefinite linear equations"; *Parallel computing*, vol. 28, pp. 323-341, 2002.

[11]Y. Liang, M. Szularz, and L. T. Yang. "Finite-element-wise domain decomposition iterative solvers with polynomial preconditioning"; *Mathematical and Computer Modelling*, vol. 58, pp. 421-437, 2013.