Ad hoc Networks Routing : Shortest Path is Enough

Maher Heal [1] **and Marwan Fayed** [2]
[1]Department of Computing Science and Math, University of Stirling, Stirling, UK,
Email: maher.heal@cs.stir.ac.uk
[2]Department of Computing Science and Math, University of Stirling, Stirling, UK,
Email: mmf@cs.stir.ac.uk

Abstract— *It is well-known in hardwired networks that shortest path routing is optimum in regard to optimizing network performance such as maximizing network throughput for instance, but little is known in this regard in ad hoc networks routing due to the dynamic, changing topology of these networks. Via linear programming formulation of the optimum routing problem, we show shortest path is the best routing strategy as well when it comes to maximizing the mean network throughput in ad hoc networks provided that interference is neglected. However, the routing metrics in selecting routes (paths) should be dependent on links availability probabilities in these networks. Heuristic approaches were used in suggesting such metrics when the links capacities are equal and unequal.*

Keywords: Ad hoc networks, shortest path routing, linear programming, routing metric, mobility

1. Introduction

Optimum routing has been studied in the context of hardwired networks. In a hardwired network where the topology is fixed and rarely changing, optimum routing can be formulated as a linear programming problem to optimize a certain objective function. It has been shown that a routing algorithm that selects the shortest path is usually the best algorithm in optimizing many objective functions such as maximizing network throughput [1].

In Ad hoc network where the topology is changing and maybe is changing randomly, little is known whether the best routing algorithm should be shortest path algorithm or not. By generalizing the linear programming formulation in [1] to model the changing topology of the ad hoc network, neglecting interference, and by assuming links between nodes are available with a constant probability, we show the best routing algorithm to maximize the mean network throughput is shortest path as well. However, the routing metric must be dependent on these links availability probabilities. Afterwards heuristic reasoning is applied to suggest ways to calculate such metrics.

2. Related Work

Many routing protocols were suggested in the literature with different routing metrics that capture different working aspects of the ad hoc environment such as link quality, interference, mobility, and energy constraints [2]. However shortest path minimum hop count is the default metric in many popular ad hoc routing protocols, such as OLSR [3], DSR [4], AODV [5] and DSDV [6]. The first experimental work that doubted shortest path minimum hop count as the right metric to achieve high throughput is that of De Couto et al. [7]. By experimenting with two testbeds of static wireless networks running DSDV protocol, they showed the protocol may select low link quality minimum-hop counted path which leads to low throughput due to retransmissions. Although in this paper we propose better metrics for wireless networks by assuming links available with constant probabilities due to changing topology, but links availability with a certain probability could be due to link qualities. In another paper De Couto et al. [8] proposed ETX routing metric based on active probing measurements and many metrics were derived and based on that metric. However, though metrics based on active measurements with probe packets are better than minimum hop count metric in static networks, the reverse is true with mobile networks [2]. Researchers suggested mobility aware metrics, for example Mcdonald and Zanti [9] suggested a routing metric that selects more stable paths based on the availability of network paths that are subject to link failures caused by node mobility. Kamal Jain et al. [10] in their seminal paper proved that shortest path is not the optimum throughput routing strategy in multi-hop wireless networks due to interference which was modeled using conflict graphs. Moreover they proved that the optimal throughput problem is NP-hard. However their analysis was for static or infrequent changed topology networks.

In this paper however, we model topology change but we neglect interference which is a limitation of more model to be incorporated in future research and extensions. However, the contribution of this paper is in suggesting and devising the mathematical means to derive better routing metrics for mobile ad hoc networks and proving shortest path is optimum in maximizing throughput whenever the interference is low and can be neglected.

3. Background

In this section we summarize the problem of optimum routing and its linear programming formulation in hardwired networks. Details can be obtained from [1].

Let us say we have a fixed topology network where the links capacities are already selected, the traffic demand is known ahead of time, the traffic is elastic[1] and we can arbitrarily split demands on different paths. Let us say also as our objective function we want to maximize the smallest spare capacity (difference between link capacity and carried load) of all links. This objective function is reasonable because by that we guarantee there will be enough capacity in the network for more traffic and hence better throughput and less delay (see the proposition in the appendix). This problem is formulated as a linear programming problem as given below.

Let the network be represented as a directed graph $\mathbb{G}(\mathbb{N}, \mathbb{L})$ where $\mathbb{N}$ is the set of nodes (routers) and $\mathbb{L}$ is the set of directed links. This means that if a and b are two nodes in $\mathbb{N}$, then the link $a \to b$ and the link $b \to a$ are distinct. Thus, any link $l \in \mathbb{L}$ has a head node and a tail node, and as the names indicate, the link is directed from the head to the tail. In ad hoc networks the case of asymmetric links is quiet common, though some routing protocols where designed to work with symmetric links only. There are K demands that are to be routed on the network. Each demand is associated with an ordered pair of nodes (n_1,n_2), where n_1,$n_2 \in \mathbb{N}$. Note n_1 is the source of the demand, and n_2 is the destination. We number the quantities as follows: demands are numbered $1, 2, ..., k, ..., K$, nodes $1, 2, ..., i, ..., N$, so that $|\mathbb{N}| = N$, and links are numbered $1, 2, ..., l, ..., L$, so that $|\mathbb{L}| = L$.

Now denote the demands by $d(k)$, $1 \leqslant k \leqslant K$, and define a flow vector $X(k)$ of dimension $L \times 1$ corresponding to the kth demand, with $X(k)_l$ represents the amount of the kth demand carried on link l, where $1 \leqslant k \leqslant K$. The topology of the network is summarized using node-link incidence matrix A of dimension $N \times L$ where

$$A_{i,l} = \begin{cases} +1 & \text{if } i \text{ is the head of link } l \\ -1 & \text{if } i \text{ is the tail of link } l \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

And define the demand vector $V(k)$, $1 \leqslant k \leqslant K$, of dimension $N \times 1$ as

$$V(k)_i = \begin{cases} d(k) & \text{if } i \text{ is the source of demand } k \\ -d(k) & \text{if } i \text{ is the destination of demand } k \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The optimization problem is given by

$$\max\ z$$

[1]Elastic traffic has no intrinsic transfer rates or end-to-end delay requirements. It is generally the traffic generated by TCP sessions like browsing.

$$\begin{bmatrix} A & \mathsf{0} & \mathsf{0} & \dots & \mathsf{0} \\ \mathsf{0} & A & \mathsf{0} & \dots & \mathsf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathsf{0} & \mathsf{0} & \dots & \mathsf{0} & A \end{bmatrix} \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(K) \end{bmatrix} = \begin{bmatrix} V(1) \\ V(2) \\ \vdots \\ V(K) \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} I & I & \dots & I \end{bmatrix} \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(K) \end{bmatrix} + z\mathsf{1} \leqslant C \quad (4)$$

$$X(k) \geq 0, 1 \leq k \leq K, z \geq 0 \quad (5)$$

where z is the minimum spare capacity. In equation (3), the left matrix is of dimension $KN \times KL$ and there are K block elements in each row and K block elements in each column; A is the node-link incidence matrix, of dimension $N \times L$. $\mathsf{0}$ is also a matrix of dimension $N \times L$. The middle matrix in equation (3) is of dimension $KL \times 1$ and the right matrix is of dimension $KN \times 1$.

In equation (4), the left matrix is of dimension $L \times KL$ and there are K block elements in it, where I is the $L \times L$ identity matrix. $\mathsf{1}$ is a column vector of L elements, all of which are 1. C is a vector of L elements where the l^{th} element is the capacity of link l.

A solution to the above linear programming problem exists when routing is done by selecting shortest paths where the link weights are the optimal dual variables of the dual problem. Coarse approximation is used in practice like taking the weights as the inverse of link capacities as in Cisco implementation of OSPF [11].

4. Generalization and Metrics

4.1 Generalization

The problem in ad hoc networks is exactly the same as the problem of hardwired networks, when mobility is the only factor of concern. The only difference is that the node-link incidence matrix A that defines the network topology is not static but dynamic with 1's and 0's entries changing according to the time. The entries (1's and 0's) could be assigned randomly depending on the mobility model used.

The optimum routing solution of this problem is unknown and it could be shortest path or not. We will show the optimization problem in ad hoc networks is the same as in hardwired networks when the mean throughput is maximized and accordingly shortest path is optimum.

Again the network is represented by a directed graph of fixed nodes and changing links from one time epoch to another. Let $\mathbb{N}$ be the set of nodes and $\mathbb{L}$ be the union of links sets at the different time epoches of the network life time. Thus between any two nodes a and b a link may be available for certain time epoch and not available for another. We will assume the link is available between any two nodes with a constant probability p. This probability is 0 for links that are available for a short period of time only in the network life time and then they disappear, although they $\in \mathbb{L}$. They have no effect on our analysis as we are concerned with the mean throughput. Assuming links are available with a constant probability is only an approximation, depending on the mobility model.

The node-link incidence matrix is a function of time now and given by

$$A_{i,l}(t) = \begin{cases} +1 & \text{if } i \text{ is the head of link } l \text{ at time } t \\ -1 & \text{if } i \text{ is the tail of link } l \text{ at time } t \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$A(t)$ is of dimension $N \times L$. The flow vector for demand k is also a function of time $X(k)(t)$; we have

$$A_i(t).X(k)(t) = \begin{cases} d(k) & \text{if } i \text{ is the source of demand } k \text{ at time } t \\ -d(k) & \text{if } i \text{ is the destination of demand } k \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

Note that each component of $X(k)(t)$ for each link, i.e. $X(k)(t)_1$, $X(k)(t)_2$,..., $X(k)(t)_L$, may be zero when the link is not available between the nodes. Now

$$A(t).X(k)(t) = V(k)$$

where $V(k)$ is the usual demand vector in equation (2). By considering the flows of all demands, we have

$$\begin{bmatrix} A(t) & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A(t) & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & A(t) \end{bmatrix} \begin{bmatrix} X(1)(t) \\ X(2)(t) \\ \vdots \\ X(K)(t) \end{bmatrix} = \begin{bmatrix} V(1) \\ V(2) \\ \vdots \\ V(K) \end{bmatrix} \quad (7)$$

$A(t)$ is the node-link incidence matrix in equation (6).

Let $C(t)$ be the capacity vector of the links. It is a vector of dimension $L \times 1$ and is a function of time. Note that $C_i(t)$ may assume the value zero when the link i is not available in certain time epoches or a constant capacity C_i, when the link is available. Then

$$X(1)(t) + X(2)(t) + \dots + X(K)(t) \leqslant C(t)$$

It is a vector inequality, where both sides are vectors of dimension $L \times 1$.
The spare capacity vector is given by

$$Z(t) = C(t) - (X(1)(t) + X(2)(t) + \dots + X(K)(t))$$

So far the analysis done is the same as that in [1], but time is added as a parameter to reflect the dynamic changing topology of ad hoc networks. Since we are interested in optimizing the mean performance of the network (mean throughput), we will try finding solutions that maximize the mean of the spare capacities of links. Taking the mean of both sides of the above equation, the mean spare capacity is given by

$$E(Z(t)) = E(C(t)) - (E(X(1)(t)) + E(X(2)(t)) + \dots + E(X(K)(t))) \quad (8)$$

Note

$$E(C(t)) = \begin{bmatrix} p_1 C_1 \\ p_2 C_2 \\ \vdots \\ p_L C_L \end{bmatrix} \quad (9)$$

where p_i is the probability that link i is available.

In hardwired networks, we want to maximize the spare capacity of links and thus we maximize the minimum spare capacity since by that we guarantee all links will have a spare capacity more than that minimum which is maximized. In ad hoc networks we want to maximize the mean spare capacity of all links and hence we will maximize the minimum mean spare capacity[2] given by equation (8).

Let $z = min_{l \in \mathbb{L}} E(Z(t))$, then we have

$$E(X(1)(t)) + E(X(2)(t)) + \dots + E(X(K)(t)) \leqslant E(C(t)) - z\mathbf{1}$$

where $\mathbf{1}$ is a column vector of L elements, all of which are 1. In matrix form

$$\begin{bmatrix} I & I & \dots & I \end{bmatrix} \begin{bmatrix} E(X(1)(t)) \\ E(X(2)(t)) \\ \vdots \\ E(X(K)(t)) \end{bmatrix} + z\mathbf{1} \leqslant \begin{bmatrix} p_1 C_1 \\ p_2 C_2 \\ \vdots \\ p_L C_L \end{bmatrix} \quad (10)$$

I is the $L \times L$ identity matrix and there are K blocks in the left matrix of the above equation and thus it is of $L \times KL$ dimension. By taking the mean of both sides of equation (7), our optimization problem is

$$\max z$$

[2]Although the mean spare capacity of each link will be guaranteed to be more than that minimum, but the minimum mean spare capacity could be due to a link that is available for a short period of time, i.e. with very low probability and hence we could maximize the minimum of mean spare capacity divided by availability probability. In that case I in left matrix of equation (10) must be the identity matrix where its diagonal 1's replaced by $1/p_i$,$1 \leqslant i \leqslant L$. Also the right matrix is replaced by links capacities vector.

$$E\left(\begin{bmatrix} A(t) & 0 & 0 & \dots & 0 \\ 0 & A(t) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & A(t) \end{bmatrix}\begin{bmatrix} X(1)(t) \\ X(2)(t) \\ \vdots \\ X(K)(t) \end{bmatrix}\right) = E\left(\begin{bmatrix} V(1) \\ V(2) \\ \vdots \\ V(K) \end{bmatrix}\right) \quad (11)$$

$$\begin{bmatrix} I & I & \dots & I \end{bmatrix}\begin{bmatrix} E(X(1)(t)) \\ E(X(2)(t)) \\ \vdots \\ E(X(K)(t)) \end{bmatrix} + z\mathbf{1} \leqslant E(C(t)) \quad (12)$$

$$E(X(k)(t)) \geq 0, 1 \leq k \leq K, z \geq 0 \quad (13)$$

To find the mean of left side of equation (11), consider $A(t).X(k)(t)$; it is clear the ith entry of this multiplication is given by

$$\sum_{j=1}^{L} a_{ij}(t).X(k)(t)_j$$

where

$$a_{ij}(t) = \begin{cases} +1 & \text{if } i \text{ is the head of link } j \text{ at } t \\ -1 & \text{if } i \text{ is the tail of link } j \text{ at } t \\ 0 & \text{otherwise} \end{cases}$$

$a_{ij}(t) = 1$ or -1 with probability p_j and zero with probability $1 - p_j$, and whenever $a_{ij}(t)$ is zero $X(k)(t)_j$ is zero as well. Hence, we have:-

$$E(a_{ij}(t)X(k)(t)_j) = \begin{cases} +E(X(k)(t)_j) & \text{if } i \text{ was the head of link } j \\ -E(X(k)(t)_j) & \text{if } i \text{ was the tail of link } j \\ 0 & \text{otherwise} \end{cases}$$

Hence, equation (11) can be written as

$$\begin{bmatrix} A & 0 & 0 & \dots & 0 \\ 0 & A & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & A \end{bmatrix}\begin{bmatrix} E(X(1)(t)) \\ E(X(2)(t)) \\ \vdots \\ E(X(K)(t)) \end{bmatrix} = \begin{bmatrix} E(V(1) \\ E(V(2)) \\ \vdots \\ E(V(K)) \end{bmatrix} \quad (14)$$

$$A_{i,l} = \begin{cases} +1 & \text{if } i \text{ is the head of link } l \\ -1 & \text{if } i \text{ is the tail of link } l \\ 0 & \text{otherwise} \end{cases}$$

By this we have proved the optimization problem in ad hoc networks is the same as hardwired networks but with with variables $E(X(k)(t))$'s and z which is now the minimum mean spare capacity. Also the link capacities now are multiplied by the links availability probabilities. Since the problem is the same, then shortest path is expected to be the optimum routing in ad hoc networks but the routing metric needs to be estimated based on link availability probabilities.

4.2 Metrics

In our analysis we assumed the probability of node A being within range of node B in ad hoc network is p. Then the link between A and B will be available p of the time and unavailable $1 - p$ of the time, i.e. The channel capacity of the link between A and B is pC on average; where C is the link capacity of the link between A and B when they are within range of each other. This is similar to time division multiplexing between two connected nodes in hardwired networking where the capacity is the proportion of the time allocated for communication between A and B; though it could be better approximated by statistical multiplexing.

By applying shortest path algorithm and using the inverse of capacity as links weights as in Cisco implementation of OSPF [11], the route (path) metric will be

$$\text{Metric} = 1/p_1C_1 + \cdots + 1/p_iC_i + \cdots + 1/p_MC_k \quad (15)$$

where p_i is the availability probability of the link i of the path (the path has M links; 1 to M).

When all links capacities are equal which is not usually the case in ad hoc networks, then the metric is simply the addition of the inverses of links probabilities.

The case of symmetric links metric can be derived using a different approach. Rather than trying to maximize mean spare capacity to maximize mean throughput, we will try to minimize the delay and hence the throughput is maximized. Assume retransmissions between any nodes is infinite and given the propagation delay is small compared to the transmission delay, the overall end-to-end delay of the packet is directly proportional to the total number of hop transmissions (including retransmissions). Hence the mean delay is

$$1/p_1 + 1/p_2 + \cdots + 1/p_M$$

on a path of M links; where p_i is the probability that the link i is available.

That can be easily seen by noting that the packet should pass link by link until reaching the destination at the end of link M and that the probability of passing link i is a geometric random variable with parameter p_i.

By comparing the above metric with the metric in equation (15) for the general case of different links capacities network, we see it is only the special case of that metric when all the links are of equal capacity.

5. Conclusion and Future Work

Shortest path routing has been proved as the optimum strategy (algorithm) in regard to optimizing mean network throughput in ad hoc networks whenever interference is neglected. That was done by showing the optimization problem has the same form as in hardwired counterpart through a generalization of the hardwired networks optimization problem to include the dynamic topology of ad hoc networks and assuming nodes are within range of each other with constant probabilities. However, the routing metrics must be dependent on these probabilities since they are parameters of the optimization problem of ad hoc networks. Heuristics were used to suggest such metrics.

More accurate metrics need to be estimated based on solutions of the optimization problem of ad hoc networks. Simulation studies to validate these metrics or the ones suggested using heuristics in optimizing network performance are required as well, by applying the new metrics to some of the popular routing protocols such as DSDV, DSR, OLSR or AODV. Ways to calculate such metrics by ad hoc hoc networks nodes and ways of realizing them in a distributed routing algorithm is an open problem.

The assumption of nodes are available with constant probabilities within each other range needs to be studied further depending on the mobility models. we modeled a network with fixed nodes and the topology is changing due to links availability only. A more general model is when nodes can join and leave randomly, and thus the topology is changing due to varying number of nodes as well. Finally, as interference is an important factor in deteriorating wireless networks performance, an extension of our model by incorporating interference as an additional constraint in the optimization problem is required.

Appendix I
Proposition: Proof of any Solution of Load (Throughput) Maximization is also a Solution of Spare Capacity Maximization

Let us name a routing strategy (algorithm) that solves our optimization problem, i.e. maximizing the minimum spare capacity γ and other strategies that don't solve our optimization problem ω.

In any network: (i) when γ is used, which means the minimum spare capacity is maximum, then we can inject more load (traffic) in the network, i.e. we can increase the demand (load), if it is not maximum. This means higher throughput. (ii) On the other hand when we have the demand (load) maximum then the routing strategy (algorithm) used must also be a solution to our optimization problem (maximizing the minimum spare capacity) and the maximum minimum spare capacity is zero.

Proof:

The first part of the proposition (i) is clear because spare capacities of all links are at least $z = min_{l \in \mathbb{L}} z_l$, z_l is the spare capacity of link l, $1 \leq l \leq L$ and that minimum is at maximum value when γ strategy (algorithm) is used. Hence we can increase the load of each link by that minimum which means we can increase the load (demand). We prove the 2nd part of the proposition (ii) as follows:

Let the load is maximum whatever the routing strategy used ω or γ, i.e. for the collection of all routing strategies (algorithms), then $min\ z_l$ should be zero for at least one value of l, $1 \leq l \leq L$. Let this is not the case, i.e. z_l is not zero for all values of l. Take now any path from a source node to a destination node. Increase the traffic (flow) on the links of this path by $min_{l \in \text{links of the path}} z_l$ and thus we were able to increase the total load (demand). This is a contradiction because the load is maximum and accordingly $min\ z_l$ should be zero for at least one value of l, $1 \leq l \leq L$. We have proved for any routing strategy γ or ω that at least one spare capacity is zero. Thus $z = min_{l \in \mathbb{L}} z_l = 0$ is independent of the routing strategy. Thus $max\ z = 0$ for the collection of all routing strategies (algorithms) ω or γ. Now when we use γ, we get z maximized and when we use any ω strategy (algorithm) we have z less than its maximum value but this means $z < 0$ which is a contradiction since $z \geq 0$. Hence only strategy γ can be used when the load is maximum.
Q.E.D

References

[1] A. Kumar, D. Manjunath, and J. Kuri, *Communication Networking-An Analytical Approach*. San Francisco, CA: Morgan Kaufann Publishers, 2004, ch. Shortest Path Routing of Elastic Aggregates, pp. 677–711.

[2] G. Parissidis, M. Karaliopoulos, R. Baumann, T. Spyropoulos, and B. Plattner, "Routing metrics for wireless mesh networks," in *Guide to Wireless Mesh Networks*, S. Misra, S. C. Misra, and I. Woungang, Eds. Springer-Verlag, London, 2009, pp. 199–230.

[3] T. Clausen and P. Jacquet, "Optimized link state routing protocol," IETF RFC 3626, October 2003.

[4] D. B. Johnson, D. A. Maltz, and J. Broch, "DSR : The dynamic source routing protocol for multihop wireless ad hoc networks," in *Ad Hoc Networking*, C. E. Perkins, Ed. Addison-Wesley, 2001, pp. 139–172.

[5] C. E. Perkins and E. M. Royer, "The ad hoc on-demand distance-vector protocol," in *Ad Hoc Networking*, C. E. Perkins, Ed. Addison-Wesley, 2001, pp. 173–219.

[6] C. E. Perkins and P. Bhagwat, "DSDV : Routing over a multihop wireless network computers," in *Ad Hoc Networking*, C. E. Perkins, Ed. Addison-Wesley, 2001, pp. 53–74.

[7] D. S. J. De Couto, D. Aguayo, B. A. Chambers, and R. Morris, "Performance of multihop wireless networks: shortest path is not enough," *Computer Communication Review*, vol. 33, no. 1, pp. 83–88, 2003.

[8] D. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *Proc. ACM Mobicom*, San Diego, CA, USA, Sept. 2003.

[9] A. Mcdonald and T. Znati, "A path availability model for wireless ad-hoc networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, New Orlean, LA, USA, Sept. 1999.

[10] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," in *Proc. ACM Mobicom*, San Diego, CA, USA, Sept. 2003, pp. 66–80.
[11] W. Odom, *CCNA ICND2-Official Exam Certification Guide*. Indianapolis, IN: Cisco Press, 2007, ch. OSPF, pp. 343–375.