Effect of a Label on Items for Their Popularity

Yuki Sonoda[1] and Daisuke Ikeda[2]

Department of Informatics, Kyushu University
744 Moto-oka, Fukuoka 819-0395, Japan.
[1]yuki.sonoda@inf.kyushu-u.ac.jp
[2]daisuke@inf.kyushu-u.ac.jp

Abstract— *We study popularity dissemination on items, such as products. Popularity is characterized by extreme imbalances since it is a typical rich-get-richer phenomenon. Existing researches focused on effect of consumers to dissemination, but not on effect of target items. Inspired by the result in [1], we hypothesize that some structure, called a* label*, on items increases imbalances of popularity. A category of products attached by a firm is a directly attached label while consumers can put a label on items. The goal of this paper is to confirm effect of a label on items to information dissemination. To this end, we have conducted multi-agent simulations about a virtual market in which firms produce items and consumers buy some of them. Comparing sales figures of items with labels and those without labels, we have confirmed that labels cause imbalances of popularity.*

Keywords: Multi-agent simulation, Consumer behavior, Power law distribution

1. Introduction

In this paper, we consider popularity of various things, such as smash hits of songs or movies and bursty words of blog entries, from the prospect of information dissemination. Counting target objects, we call *items*, in various fields, the common phenomenon, the power law distribution for popularity, emerges [2], [3].

In general, consumers have a lot of chances to see popular items and thus we see rich-get-richer phenomena [4], [5]. Moreover, consumers utilize information from many kinds of communication, such as word-of-mouth communication. As a result, such communications cause extreme imbalances and lead to the power law distribution. In this sense, we can think that popularity is caused by information dissemination [2]. From the view point of effective dissemination, existing researches focused on users, their network structures, and interactions of them [6], [7], [8], while the structure of target items has been ignored.

Generally speaking, however, an extremely popular trend does include many items. For example, there exists a trend of "premium" products in Japan, where a smash hit of a high-quality, high-price product yields this trend and then conversely this trend gathers similar high-quality, high-price products among different categories. We show another example of "Yuru-chara", which is

> a Japanese term for a category of mascot characters; usually created to promote a place or region, event, organization or business,

according to this page[1]. This concept became very popular after some popular characters, such as Hikonyan and Kumamon, became famous, and then the popularity of the concept created a huge number of new characters.

We can think that such a trend is a kind of categories of items. We call it a *label* of items and expect that a popular label can create an extremely large hit phenomenon. This idea was inspired by an earlier result of the authors [1], which is a research to predict potentially popular hash-tags, that is labels, of a micro-blog service. In this research, we experimentally showed that, for any label, the ratio of the number of items used with the label to the number of different items with the same label is constant over time. In other words, if the ratio is high for a label, we see many different items with the label as the label is used. Using this interesting property, we proposed a method to find latent popular labels. Although the proposed method works fine, we can not elucidate why such a phenomenon happens for labels and items.

As a first step to elucidate the mechanism of labels and items, we try to show that we see more smash hits if we can use labels. That is, the goal of this paper is to verify the following hypothesis: a label on items can increase imbalances of popularity of them. A typical label, for example, is a category of news articles while we can think anything we can attach to one or more items as a label, such as a catchy copy, an attribute, and a hash-tag.

To achieve this goal, we utilize multi-agent simulations, where we prepare consumer agents and firm ones, consumer agents can attach a label on items, both types of agents can recognize popular labels, and firm agents tends to create a new product with a popular label. Calculating sales figures in this virtual market, we compare sales of items with and without labels, and verify the effect of labels.

This paper is organized as follows: After reviewing related work in Section 2, we will explain our assumption about

[1]https://en.wikipedia.org/wiki/Yuru-chara

labels and the model for simulation in Section 3. Then we will show experimental results in Section 4. Finally, we conclude in Section 5.

2. Related Work

Popularity of items has been extensively studied in terms of information dissemination.

Some researches focused on consumers and proposed types of consumers, according to their behaviors. In [6], diffusion of innovations were studied and consumers were categorized, such as innovators and early adopters. Similary, the concept of connectors, mavens, and salesmen was proposed in [8].

There exists researches that studied dynamics of changes in popularity, basically using differential equations. For example, the Bass model was proposed in [7], considering innovation and imitation effects of consumers. Similarly, in [9], both direct communication and indirect one, such as the rummor effect, were considered and shown that the proposed equation well describes hit phenomena of movies.

Some other researches, such as [10], focused on the network structure of consumers.

However, these existing researches studied effects of communication among consumers, but not effects of structures among target items, w.r.t information dissemination. Of course, practitioners in marketing segments must know about effects and impacts of catchy slogans and copies, which can be seen as labels, since we see a lot of them on the media. However, there does not exist a quantitative research on effect of labels w.r.t information dissemination as far as the authors know.

3. Methods

In this section, we first explain, using examples, our assumption on labels for consumer behaviors, and then show our model of simulations.

3.1 Assumption of Effect on Labels

Consider that there exists a label attached with some items (see Fig. 1). Originally, the main target of a consumer is items and a label is just an attachment. For example, consider mascot characters, called "Yuru-chara." Although this term was coined in the early 2000's, the term was not so popular initially. In this case, the term "Yuru-chara" is just an attribute of mascot characters (the top figure of Fig. 1). However, after appearance of some attractive characters, such as Kumamon and Hikonyan, which are left- and right-hand side chracters in Fig. 1, the popularity of the term is risen sharply. Then we see this term many times on the media and many new characters were created. Now consumers first recognize the label on items (the bottom figure of Fig. 1). This is a *reverse phenomenon* since a popular label itself has its own popularity and increases popularity of items with the label.

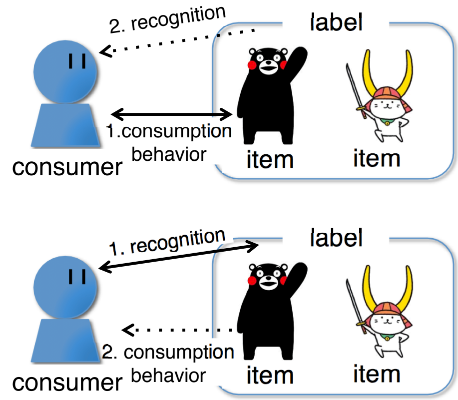

Fig. 1: Two processes of consumption behaviors of items and recognision of labels on them

One import property of labels is that the same label can be used in different fields of products. In other words, firms of other types of products can receive a label as a message. Then they can produce other types of products with the same label. An example of this phenomenon, we use "The Premium Malt's"[2]. From its name, outlook, and price, many consumers recognize "premium" as a label for it. The product was a blockbuster in Japan and made "premium" label as a popular label. In fact, we saw many high-quality, high-price products, such as premium pet food, premium canned coffee, and premium bananas, after this product. Therefore consumers have a lot of chances to see a popular label and so the popularity of the label could be drastically increasing. Then some consumers choose some products because they have the popular label.

We should note that hit products do not always create popular labels. In fact, a similar high-quality and blockbuster beer "Yebis"[3], which started to sell long before "The Premium Malt's" and a long-seller product. But it did not cause similar labels.

From the above observation, we hypothesize that a popular label on items encourages hit products and this causes imbalances of popularity among items.

3.2 Our Model

There exists two types of agents, consumer and firm agents. A consumer agent has perceived recognition rate for each label [11], choose items based on information, including the rate and rumors, and exchange information about items with other agents, where each consumer agent has randomly decided receiving and sending rates. A firm agent receives information about consumers' recognition of labels via market researches, recognize labels of product

[2]http://the-premium.jp/pc/index.html
[3]http://www.sapporobeer.jp/yebisu/

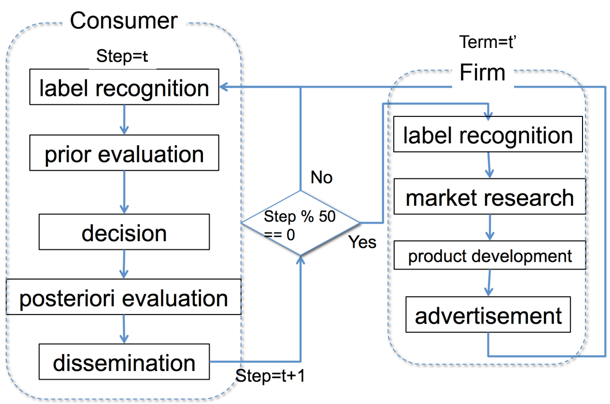

Fig. 2: The flow chart of our simulation

created by other firms, development a new product based on the sales of the last year.

The recognition of labels of other firms' product can be seen as information exchange among firms. In a simulation with labels, popularity of labels is important to develop new products, that is, a firm is likely to create a product with a popular label.

Fig. 2 shows the flow chart of consumer and firm agents. Based on AISAS model, which is a hierarchy model of advertisements and says consumer behavior changes in order attention, interest, search, action, and share [12], we decide one step of consumer agents as follows: they first recognize labels, then evaluate items before purchase, then choose an item, then evaluate purchased item, and finally disseminate information about the item. Similarly, we decide one step of firm agents as follows: they first recognize labels, then conduct market researches, then develop new products, and finally advertise them. In simulations without labels, both types of agent ignore tasks related to labels.

Every step of the firm agents is executed after 50 steps of the consumer agents because one step of firm agents takes longer time, compared to daily consumer behavior. For simplicity, one firm creates only one product and thus if a firm creates a new product, the current product of the firm is removed.

4. Results

After explaining environments for our experiments, we show results of our experiments, including preliminary ones for choosing some parameters, such as the number of trials.

4.1 Environments

Based on the model described in the previous section, we implemented the simulation program in Python and compiled with Cython. All experiments were executed on MacBook Pro (OS:Mac OS X 10.8.5, CPU:2.9GHz Intel Core i7, Memory:8GB 1600MHz DDR3).

Table 1: Parameters for our simulation

parameter	its value
# consumers	900
# firms	100
# types of labels	50

Table 1 shows the values for some parameters used in our simulation program. The initial values for the following parameters are real values randomly decided in $[0, 1]$: preference of consumers, sending and receiving moods of word-of-mouth, and advertisements of firms.

To evaluate results of simulations, we basically use histograms of sales among all firms since popularity of products are known to follow the power law distribution.

4.2 Preliminary Experiments

In this section, we show two results of preliminary experiments and decide the values for the following two parameters: the number of trials for stable results and the number of types of labels. To reduce the times required by preliminary experiments, we set the number of consumers to be 100 while it is 900 in the main experiments.

As described above, random numbers are used for some parameters. To obtain stable results, we should create histograms after several trials of our simulation program. First, to decide the number of trials, we compare two histograms of sales for 30 and 100 trials.

Fig. 3 shows two histograms: red one is from 30 trials and blue one 100. The horizontal axis is bins of sales, where

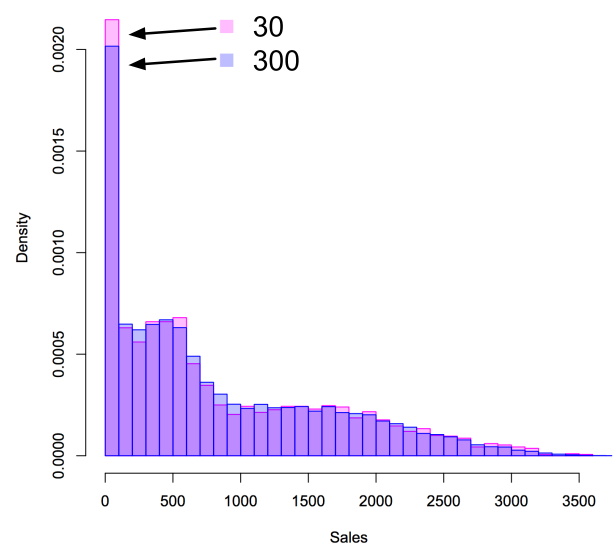

Fig. 3: Two histograms of sales for 30 and 100 trials, red one is from 30 trials and blue one 100.

the width of each bin is 100, and the vertical axis is the probability of firms, which shows how many firms go into one bin.

From Fig. 3, we can see that 30 trials are enough stable, compared to 100 ones.

Since the goal of this paper is to show effect of labels for popularity of items, we compare simulations with and without labels. In this perspective, we need the large number of different labels. However, we expect that there exists two or more items which have the common label since the label is a category of items. In this sense, the number of different labels must be smaller than the number of items, which is equal to the number of firms in our setting.

Thus the second preliminary experiment is to decide the number of different labels for main experiments. In this experiments, we compare histograms of simulations in case that the number of different labels is 5, 20, or 50. Fig. 4 shows the three histograms, where the vertical axis shows

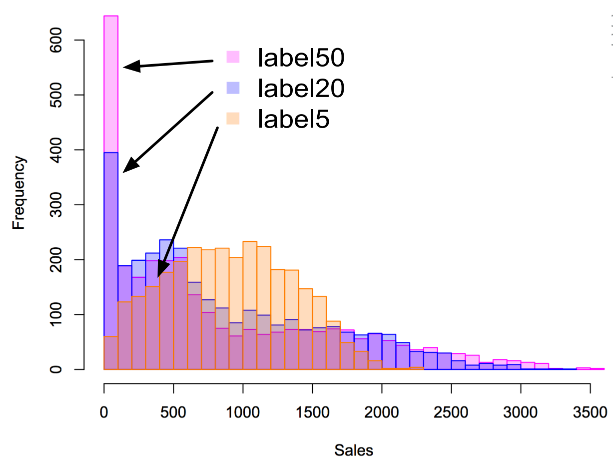

Fig. 4: Three hisgrams of sales in case that the number of types of labels is 5, 20, or 50.

the number of firms, that is the frequency, whose sales fall in the corresponding bin, the horizontal axis is bins of sales, where the width of bins is 100, and yellow (resp. blue and red) histograms are in case that the number of different labels is 5 (resp. 20 and 50).

From Fig. 4, we find that the shape of a histogram become skewed and the mode value become smaller as the number of different labels is increasing.

Table 2 shows statistics, such as averages, in case that the number of labels is 5, 20, or 50. As the number of different

Table 2: Statistics in case that the number of labels is 5, 20, or 50.

# types of labels	average	median	skewness
5	906.2183	900	0.1032524
20	849.9577	631	0.8267709
50	834.0787	549	0.9537798

labels is decreasing, values of the average and median is increasing. In this sense, the fewer the number of different labels is, the more total sales are achieved. On the other hand, the skewness become much larger as the the number of different labels is increasing. In this sense, the number of labels has impact on imbalances of popularity. Considering our goal of this paper, we use 50 as the value for the number of different labels.

4.3 Main Results

As the main results, we show three types of graphs: one is dynamics of one trial and the other two are distributions of all trials.

Fig. 5 shows transitions of sales for each label in one trial of simulation, where the horizontal axis shows the total 1500, which equals to 50 steps times 30 trials, and the vertical axis the sales. In this graph, we see that the sales of the blue label

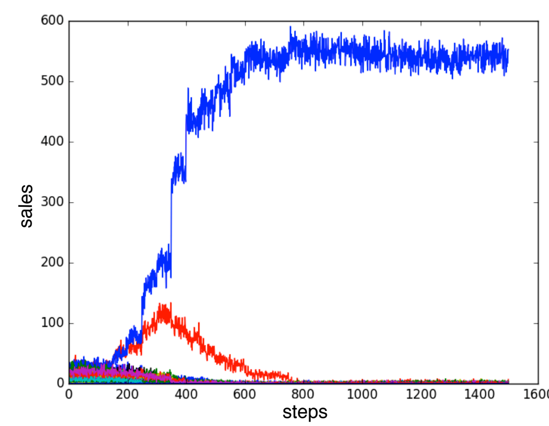

Fig. 5: Transitions of sales for each label in one trial of simulation, where each line shows the total sales of the products with the same label.

is drastically risen around 400 steps. This is a typical rich-get-richer phenomenon. We confirmed similar phenomena in all the other trials.

We created two types of graphs from 30 trials: one is rank-size plots (see Fig. 6) and the other histograms of sales (see Fig. 7).

Fig. 6 shows two rank-size plots, where the vertical axis shows sales for firms and the horizontal axis ranks of firms in decreasing order of sales. The left-hand (resp. right-hand) side graph is one created from 30 trials of simulations without (resp. with) labels.

We find that the amount of sales at the top rank in the right-hand side graph is about twice of that in the left-hand one, and the curve in the left-hand graph is smoothly declined, compared to that in the right-hand one.

Fig. 7 shows two histograms, where the horizontal axis shows sales bins, each of whose width is 300, and the vertical

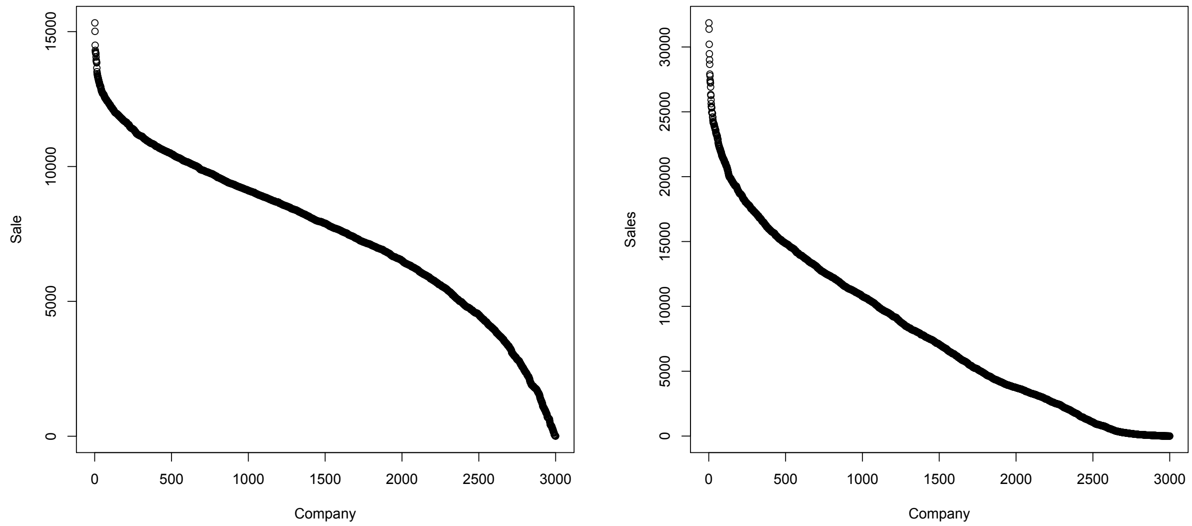

Fig. 6: Totals sales of 100 firms in 30 trials without labels (left-hand side) and those with labels (right-hand side) are plotted in decreasing order.

axis the frequency, that is, the number of firms whose sales figures is in some bin. The blue (resp. red) histogram is created from 30 trials of simulations without (resp. with) labels.

We find that the blue histogram is unimodal, where the mode is around the average value (see Table 3). In the red one, the mode is at the lowest bin of sales and there exists many firms achieving much higher sales, compared to top sales in the blue one.

From Fig. 7 and Fig. 6, we can conclude that labels cause imbalance of pupularity.

5. Conclusion

In this paper, we have considered impact of the label on items w.r.t popularity of them, conducting multi-agent simulations. We have compared two types simulations, that is, one without labels and the other with them. Although the histogram created from the simulation without labels is unimodal and the mode is at around the average, only introducing the structure into target items by labels causes imbalance of popularity of items. While existing researches about popularity of items focused on interactions and/or structures of consumers, imbalance of popularity can be achieved by the structure of items. As far as the authors knows, this is the first result which reveals that the structure of items is critical to popularity of items.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 15H02787.

References

[1] Y. Sonoda and D. Ikeda, "Predicting Latent Trends of Labels in the Social Media Using Infectious Capacity," *International Journal of Future Computer and Communication*, vol. 4, no. 6, pp. 374–380, 2015.
[2] M. E. J. Newman, "Power Laws, Pareto Distributions and Zipf's Law," *Contemporary Physics*, vol. 46, pp. 323–351, 2005.
[3] D. Sornette, *Rank-Ordering Statistics and Heavy Tails*. Springer, 2000, ch. 6 of Critical Phenomena in Natural Sciences.
[4] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, pp. 509–512, 1999.
[5] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.
[6] E. M. Rogers, *Diffusion of Innovations*, 3rd ed. Free Press, 1982.
[7] F. M. Bass, "A New Product Growth for Model Consumer Durables," *Management Science*, vol. 15, no. 5, pp. 215–227, 1969.
[8] M. Gladwell, *The Tipping Point: How Little Things Can Make a Big Difference*. Little Brown, 2000.
[9] A. Ishii, H. Arakaki, N. Matsuda, S. Umemura, T. Urushidani, N. Yamagata, and N. Yoshida, "The 'Hit' Phenomenon: a Mathematical Model of Human Dynamics Interactions as a Stochastic Process," *New Journal of Physics*, vol. 14, 2012.
[10] J. J. Brown and P. H. Reingen, "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research*, vol. 14, no. 3, pp. 350–362, 1987.
[11] H. Yamamoto, S. Nishida, S. Morioka, and S. Yamakawa, "The Effect of "Perceived Recognition Rate" on Consumer Behavior," *Journal of Marketing Science*, vol. 19, no. 1, pp. 73–89, 2011, in Japanese.
[12] R. Akiyama and K. Sugiyama, *Holistic Communication*. Sendenkaigi, 2004, in Japanese.

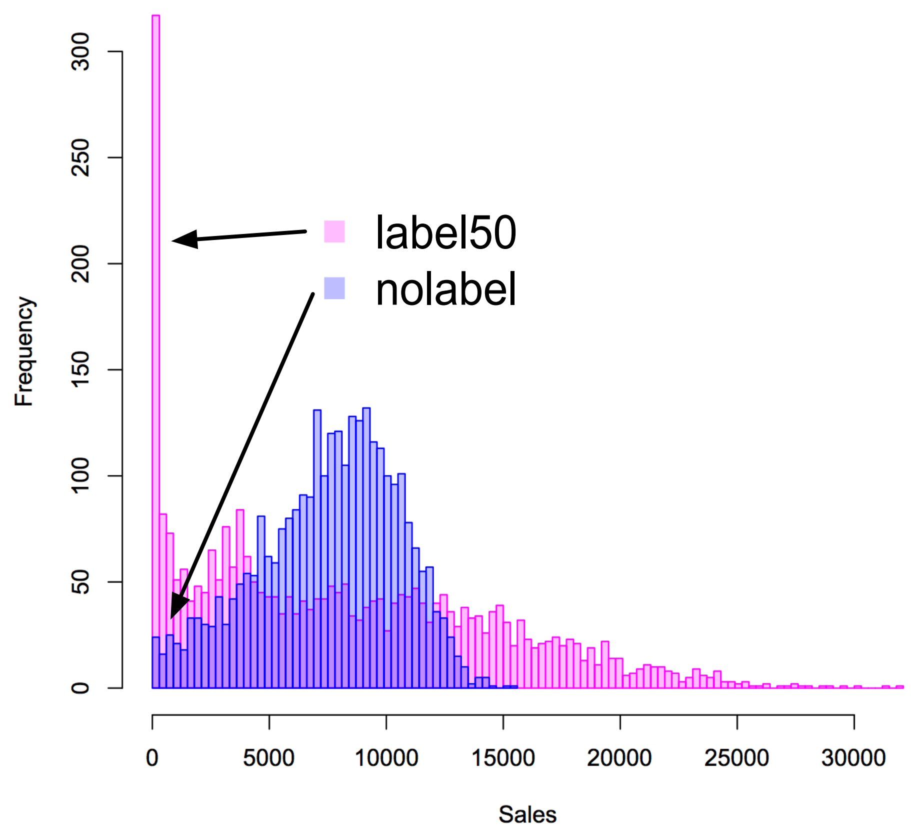

Fig. 7: Histogram of sales for simulations without labels (the blue histogram) and with labels (the red one)

Table 3: Statistics with or without labels.

	average	median	skewness
without labels	7540.16	7895	-0.4017727
with labels	8042.85	7056	0.6286133