

Student: Thanh V Nguyen
Student ID: 3534920
Creating the Social Web – Assignment 1

Goal:

Many tagging systems provide popular tags that have been previously applied to resources as hints to help taggers to choose appropriate tags when tagging. Many researches like [1] agree that this will affect personal tagging behavior. Typically, we simply think that the tags which have been used most will be the popular tags. However, it is clear that the each resource need time to get a stable set of popular tags. So showing the most used tags of resources as popular tags when there are not many people tag on them may not be good. For example, the most used tags when there are only 1 or 10 people tagging in most cases are not reliable comparing with the most used tags when there are 100 or 1000 people tagging.

So, in this experiment, I want to examine two following research questions (RQs).

RQ1: How long a most used tag should be considered as a popular tag?

RQ2: Does the time we decide to begin showing the popular tags affect the tags distribution?

The time dimension indicated in two RQs is calculated by the number of taggers tagging on that resource.

Method:

The experiment is simulated using Repast [2]. The behavior probability is approximation by using the results in [1]. We assume that there is one hottest movie in our system and all users will like to tag on it. Like [1], we separate three kinds of tagging behaviors, *invented*, *reinvented* and *borrowed*, with the corresponding probability 76%, 4% and 20%. Every step in the simulation will be considered as a new user tag on that movie. Taggers may *invent*, *reinvent* or *borrow* tags in each step. We will examine the distribution of the most used tags during the experiment.

Results:

In all of experiments, we assume total 1000 users will tag on the movie. In first experiment (Fig 1), the popular tags will be shown immediately when users tag on the movie.

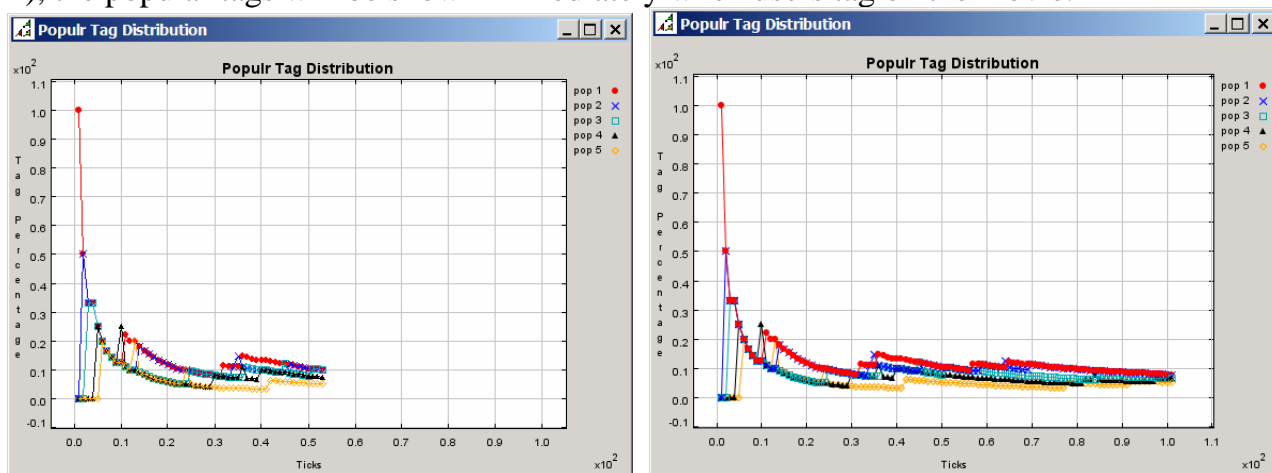


Fig 1: (a) After 50 users tagging

(b) After 100 users tagging

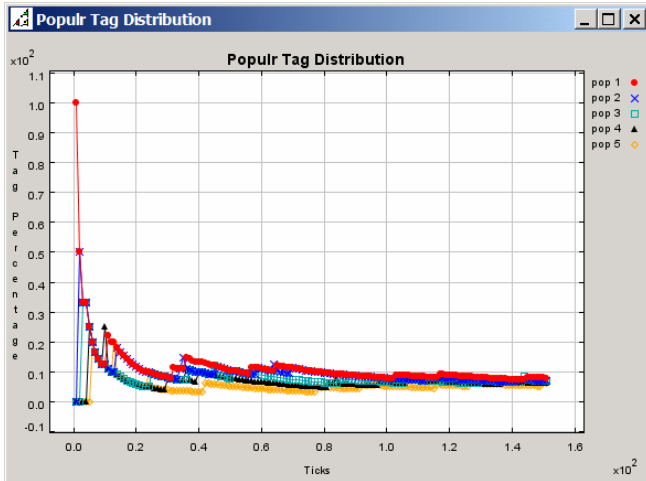
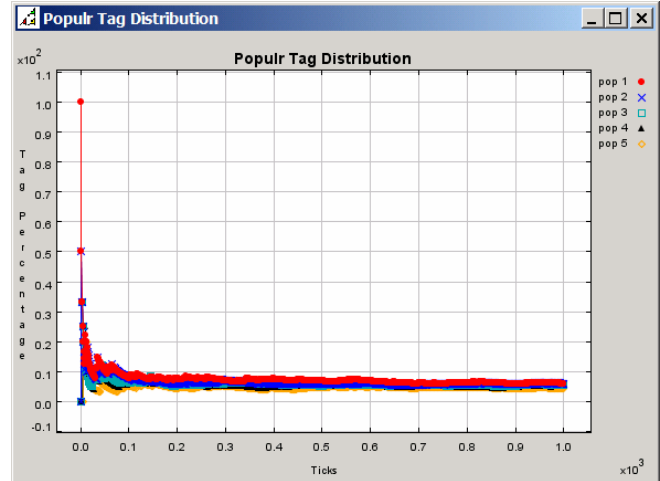


Fig 1: (c) After 150 users tagging



(d) After 1000 users tagging

In second, third and fourth experiment (Fig 2,3,4), the popular tags will be shown after at least 50, 100 and 500 users have tagged on the movie.

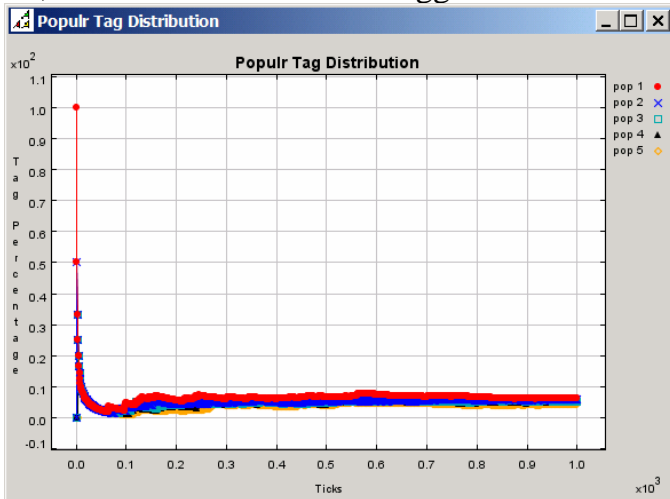


Fig 2

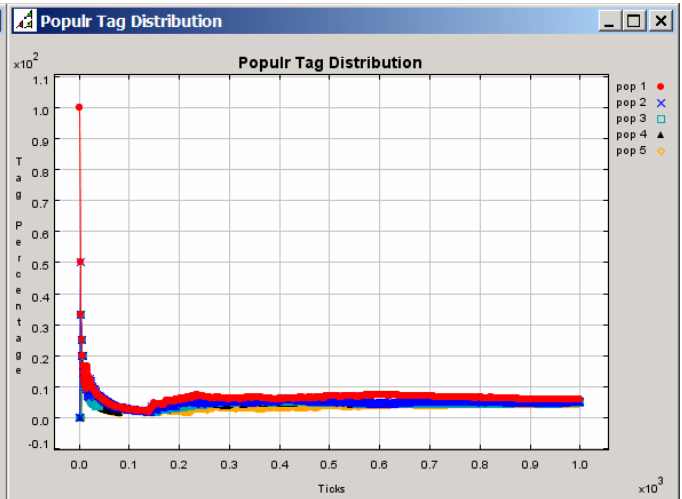


Fig 3

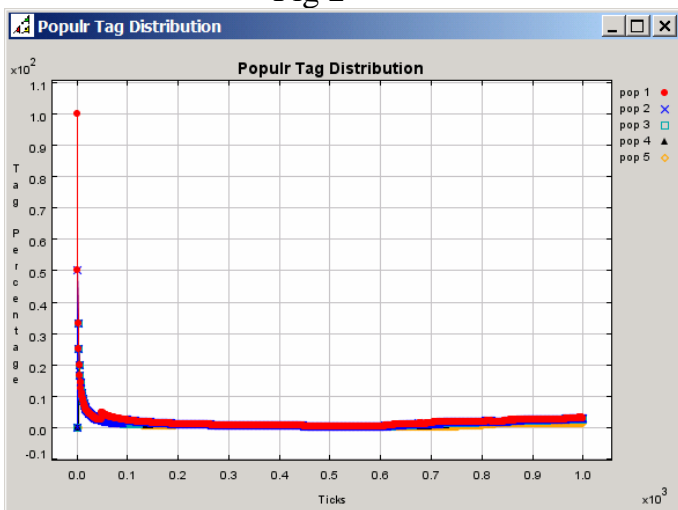


Fig 4

Discussion:

All of the figures show that the popular tags will become stable after a large number of people tag on the movie. Especially, in first, second and third experiment (Fig 1d, 2 & 3) the popular tags get the approximate percentage, about 4-5%. Now we answer to two RQs stated above.

RQ1: How long a most used tag should be considered as a popular tag?

As be shown in Fig 1, the percentage of the most used tag will become stable after about 10% users involve in tagging. In real application, we can estimate the number of users using our system. So we may estimate the time to show the popular tags to the users, not necessary immediately at the beginning.

RQ2: Does the time we decide to begin showing the popular tags affect the tags distribution?

Fig 1,2 & 3 confirm that the distribution of popular tags do not change too much if we decide to show the popular tags sooner or later. This affirms that we do not need to show the popular tags immediately when the user start tagging on the movie.

I also want to recall that we only talk about the distribution of the popular tags, not the actual popular tags. ***In conclusion, the time we decide to show the popular tags may affect the actual tags be chosen as popular tags but not the distribution of them.***

One of straightforward future work will be examining the change of actual tags be chosen as popular tags if we show the popular tags immediately or later when the user begin tagging.

References:

- [1] Shilad Sen, Shyong K. Lam, Dan Cosley, Al Mamunur Rashid, Dan Frankowski, Franklin Harper, Jeremy Osterhouse, John Riedl. tagging, community, vocabulary, evolution, *Proceedings of CSCW 2006*.
- [2] Repast, <http://repast.sourceforge.net>