Increasing Message Relevance in Social Networks via Context-based Routing

Vasiliki Pouli
University of Maryland
Institute for Systems Research
College Park, MD 20742, USA
+1 301-405-3374
vpouli@umd.edu

John S. Baras
University of Maryland
Institute for Systems Research
College Park, MD 20742, USA
+1 301-405-3374
baras@umd.edu

Anastasios Arvanitis
National Technical University of Athens
Iroon Polytechniou, Zografou
157 80 Athens, Greece
+30 210 772 1402
anarv@dblab.ntua.gr

ABSTRACT

Several applications are built around sharing information by leveraging social network connections. For example, in social buying sites like Groupon, a deal is usually forwarded to interested recipients through their social graph. A primary goal is to improve user satisfaction by maximizing the relevance of the shared message to the target audience. In this work, we address this problem by proposing a context-based routing approach that exploits both user preferences and the network structure aiming at enhancing the relevance of forwarded messages and ensuring that each message will be delivered to the most interested users.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Miscellaneous; J.4 [**Social and Behavioral Sciences**]: Sociology

General Terms

Algorithms, Experimentation

Keywords

Social networks, routing, preferences, similarity

1. INTRODUCTION

Several modern applications are built around sharing information by leveraging social network connections. For example, social buying sites like Groupon [1], could use a very effective way to promote a deal through sharing it to users and their friends via social networks such as Facebook [5], Twitter [3] etc. These kind of sites enable users to create online profiles, present themselves, write posts, share information and create connections with other users that consider as friends, or belong to same groups, or might just 'like'. However, as the number of daily deals increases, deal promotion is often considered as spam by people, especially if the deals are irrelevant to them. Thus, a primary problem is to provide effective ways to share the deal information only to the users that are expected to find this deal interesting, thereby increasing both user satisfaction and profit [12].

To enhance deal promotion you have to consider how the deal content relates to the users. Apparently, the relevance of each message with the recipient node (user) is usually context-based, i.e. it depends on several factors such as the context similarity between the deal and the node, represented as a set of dimensions. Such dimensions could be for example the thematic category, price, location, etc. If we consider, e.g. a deal in a Chinese restaurant in Washington DC, it would be wise to forward it mainly to users that are located close to Washington DC and have an interest for Chinese cuisine. In this way, it is more likely that these users will accept the deal and eventually buy it. On the other hand, flooding the deal-message to all your friend connections, including those that live far away from Washington DC (e.g. Asia) or have not expressed any interest in food deals, would probably be considered as spam.

The dependence of relevance to the message context [11] calls upon developing a dynamic routing approach. In our work, we have developed a routing scheme that considers user preferences and delivers coupons to the nodes most interested, thus ensuring that the deal will arrive to the most potential buyers.

2. RELATED WORK

Recent works [9, 8, 7] have proposed routing algorithms that use social characteristics with applications on delay tolerant, ad hoc or P2P networks. In [9] the authors propose a socially-based greedy routing algorithm for delay tolerant networks. They use an n-dimensional social profile of users and utilize the Jaccard coefficient to measure the similarity between them. Then, they apply a greedy routing algorithm by selecting the nodes that are socially closer to each other. Also, in [8] the authors proposed the Bubble rap, where they utilize nodes' centralities and communities to improve the forwarding efficiency. Further, in [7], they have shown that labeling the nodes with their affiliation and forwarding messages to nodes belonging to the same affiliation as the destination, can improve forwarding performance both in terms of delivery ratio and cost. However, although these algorithms exploit user-to-user similarities, they *(i)* do not utilize their preferences in terms of the forwarded message content and they *(ii)* do not ensure message delivery to all interested recipients. Motivated by this, in this work we propose a context-based routing approach that exploits both user preferences and the network structure aiming at maximizing relevance and potential profit by ensuring message delivery to all interested nodes.

3. CONTEXT-BASED ROUTING

In this section we will describe our similarity measure that is used for calculating node-message relevance distances. Our similarity measure consists of two parts; the first part incorporates the context similarity, whereas the second part measures the (structural) network distance. Subsequently, we will present our context-based routing algorithm that utilizes our proposed node similarity measure in order to propagate the deal message to all interested recipients.

3.1 Context Model

We will first explain the context part by describing the context model that we used. We will assume that both the deal message and a users' interests/preferences can be represented using the same context model. This context information can be modeled as a vector of attributes where each attribute represents a context dimension. In general, the domain for each context dimension might take values from an ontology or hierarchy, such as the one depicted in Figure 1. Such context dimensions could be for example the content topic (e.g. food, entertainment, traveling), the deal location, price range, etc.

We assume that each user has specified her interest through a set of keywords that might take values anywhere from the attribute domain. For example a user, as seen in fig. 1, might have specified that she likes 'Thai' cuisine and 'Action' movies in 'MD' . These preferences can be explicitly provided by each user (such as Facebook 'likes') or they might as well have been extracted using entity extraction techniques by analyzing a user' posts, comments, tweets etc. We also consider that a coupon k takes values within the same context model and its deal vector D_k could consist for example the values $D_k = \{Asian,\ MD\}$, which means that it would be a coupon concerning Asian food in a restaurant in MD. We could also define weights in the context dimensions and say for example that a user likes 80% Asian cuisine.

In order to measure the context similarity between a deal message and a user's preferences, we will calculate a cosine similarity between the user i preference vector U_i and the deal k vector D_k.

$$context\_similarity(U_i, D_k) = \frac{U_i \cdot D_k}{\|U_i\|\|D_k\|}$$

3.2 Hyperbolic Embedding

Furthermore, in order to ensure message delivery to all interested recipients, our approach greedily embeds the network into the hyperbolic space. We followed the greedy hyperbolic embedding since, as it is shown in [10], every finite, connected, undirected graph has a greedy embedding in a two-dimensional hyperbolic space, i.e., *one may achieve 100% success rate with greedy routing by assigning virtual coordinates in the hyperbolic plane rather than the Euclidean plane.* Based on this embedding, each node will always have a neighbor closer to the destination and a message will never get stuck into local minima as in Euclidean space. For embedding the network into the hyperbolic space we applied the algorithm that is described in [4]. Following this approach, we constructed a spanning tree from the original graph, and assigned to each node hyperbolic coordinates from the set $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ in accordance to their parents coordinates. A greedy embedding of the spanning tree is also a greedy embedding of the graph.

We consider as network distance $network\_dist(z_1, z_2)$ between two nodes z_1 and z_2 the hyperbolic distance between them.

Figure 2 shows a random graph of 20 nodes and Figure 3 its minimum spanning tree hyperbolic embedding in the Poincaré half plane disc.

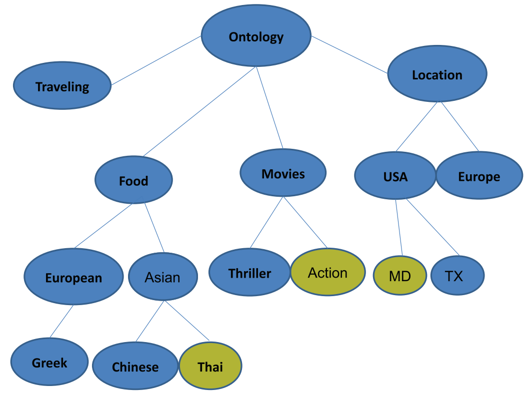

Figure 1: context ontology

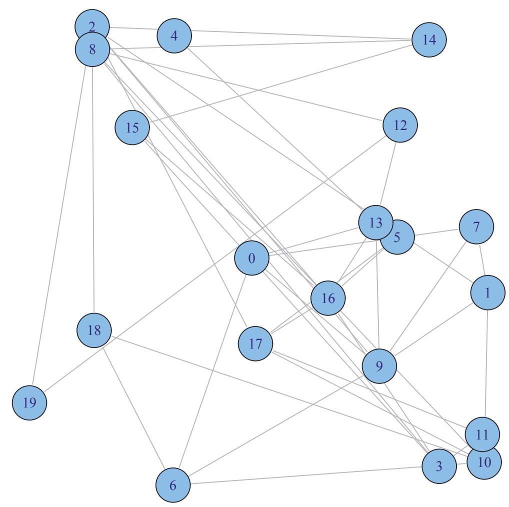

Figure 2: random graph of 20 nodes

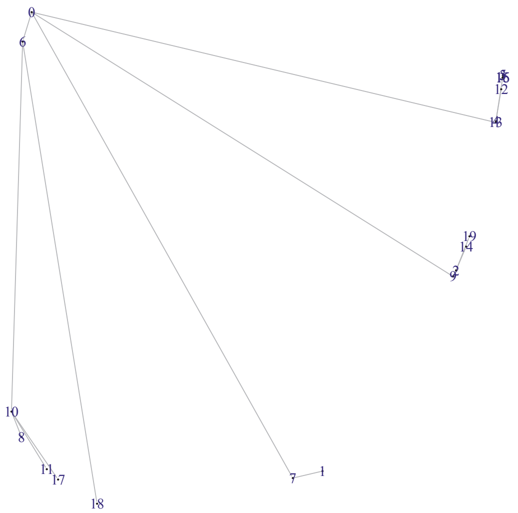

Figure 3: Greedy Hyperbolic embedding in the half plane Poincaré disc model of minimum spanning tree of graph depicted in Figure 2

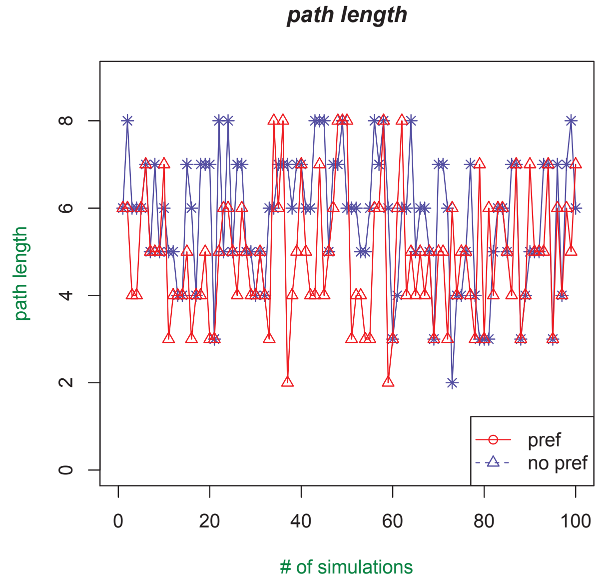

Figure 4: path length

3.3 Routing Algorithm

The goal of the routing algorithm is to ensure that each message will reach, in the minimum number of steps, the nodes that are most relevant to a deal. We will assume that the deal vendor would like to forward the message through his friends/followers' connections to the top-m most interested users of the network in the minimum numbers of steps. Based on the proposed context similarity measure and hyperbolic embedding mapping, we defined a relevance metric, named: *relevance*, given by (1). This metric incorporates both the context similarity and the (structural) network distance and calculates which node i is most relevant to a deal k:

$$\operatorname*{argmax}_{i} relevance(k,i) = a * \frac{context\_similarity(U_i, D_k)}{network\_dist(z_k, z_i)} \tag{1}$$

where a is a normalization factor that determines the relative importance that we assign to the network distance and context similarity respectively.

In order to discover these nodes, we can simply evaluate (1) on the nodes of the network using our context similarity measure and the network distance which is available using our hyperbolic graph embedding. The target nodes i.e. those that are potentially more interested to the deal., are the ones that maximize our similarity metric: $relevance(k,i)$.

After identifying the target nodes, we proceed to our routing algorithm. Initially, the algorithm begins from D, i.e., the node that issues the deal message. For example in Facebook, D might be the account of a deal provider. Next, at each step the algorithm utilizes our network embedding to find out which of our neighbors are closer to the target user than the current node. From these nodes, the algorithm selects to forward the message to the n neighbors that have the largest context similarity with the deal message. The algorithm terminates when the destination has been reached.

4. EXPERIMENTAL EVALUATION

We carried out some simulations to examine the efficiency of our algorithm against a non social-based scheme. More specifically, we used the R Project [2] and generated a undirected random network with N=100 nodes and probability for drawing an edge between two arbitrary vertices equal to 0.02. We chose m=1 meaning that we take one destination i.e. the node i that maximizes $relevance(D,i)$. Further, we also took n=1, i.e., our routing algorithm will pass the message to the one neighbor that maximizes his context similarity with the coupon and of course has smaller network distance to the destination node. Finally, we considered α=1 in (1) and an ontology model with 20 context dimensions.

In each experiment we selected random user preference and deal values for the context dimension categories and we ran the simulations 100 times by selecting a different source node for every iteration.

Through our simulations we want to compute the number of hops needed to reach the top-m destination nodes and also the path preference *pathPref* which we define as the average context similarity of the coupon with all the nodes along the path:

$$pathPref = \sum_{i=1}^{M} \frac{context\_sim(D,i)}{M}$$

where M is the number of hops along the path.

The motivation for using path preference as a metric in our experiments, is that we want to evaluate the relevance of the message with the users that belong to the path followed. The closer you are connected to a dealer node the more probable it is to like the deals he offers and the probability of buying those deals increases. Thus, path preference metric could be considered as a profit measure.

We first calculate the path length from the source to the destination using our proposed algorithm, and then we compare it against a greedy forwarding routing which selects to

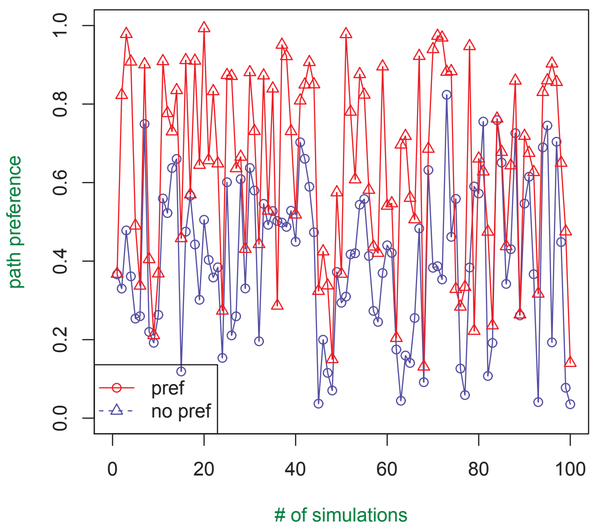

Figure 5: path preference

Table 1: Comparison of social and non-social routing algorithms

	Social vs Non-social
Path length	-86%
Path preference	+96%

forward the coupon to the neighbor closest to the destination but without taking into account the preferences of the nodes.

According to our simulations, in Figure 4 we see that in our context-aware algorithm the number of hops from the sender to the destination is decreased by 86%, compared to the non-social greedy case. This means that we get an improvement in finding the most relevant nodes in fewer steps compared to the non-social-based routing.

Second, we observe in Figure 5, that by exploiting the user preferences in our algorithm, then the whole path preference increases by 96%, compared to the simple greedy case, which does not take into account the similarity of the user with the deal.

5. A SOCIAL ROUTING APPLICATION

Our context-based social routing algorithm can be implemented as an application that runs on top of a real social network such as Facebook [6]. Using such an application, a deal provider that wants to send a deal could specify under which categories and sub-categories the deal belongs to, e.g., Asian food, thriller movies, traveling to Hawaii etc. The subscribers could as well have a similar menu that would allow them to select for which categories they would like to receive messages/deals. Thereby, the system will receive each coupon and forward it to the recipients that have shown interest to the particular categories associated with the deal.

6. CONCLUSION

In this paper we proposed a context-aware routing scheme that targets to increase the relevance of messages shared across a social network. It achieves this by forwarding the message to the most relevant nodes, taking into account both user preferences and the network structure. Our preliminary experiments show that our context-aware routing scheme provides significant improvement both in terms of the number of hops required to forward each message and the preference similarity of nodes along the path that is followed, thus increasing the chances of users that are most related to the message context to buy a deal coupon.

7. ACKNOWLEDGEMENTS

Research partially supported by the US Air Force Office of Scientific Research MURI grants FA9550-10-1-0573 and FA9550-09-1-0538.

8. REFERENCES

[1] The Groupon deal site. http://www.groupon.com/.

[2] The R Project for Statistical Computing. http://www.r-project.org/.

[3] I. Anger and C. Kittl. Measuring influence on twitter. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, i-KNOW '11, pages 31:1–31:4, New York, NY, USA, 2011. ACM.

[4] A. Cvetkovski and M. Crovella. Hyperbolic embedding and routing for dynamic graphs. In *INFOCOM*, pages 1647–1655, 2009.

[5] N. B. Ellison, C. Steinfield, and C. Lampe. The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168, 2007.

[6] J. Feiler. *How to do everything : Facebook applications*. McGraw-Hill, New York, 2008.

[7] P. Hui and J. Crowcroft. How small labels create big improvements. In *PerCom Workshops*, pages 65–70, 2007.

[8] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: Social-based forwarding in delay-tolerant networks. *IEEE Transactions in Mobile Computing*, 10(11):1576–1589, 2011.

[9] K. Jahanbakhsh, G. Shoja, and V. King. Social-greedy: a socially-based greedy routing algorithm for delay tolerant networks. In *MobiOpp*, pages 159–162, 2010.

[10] R. Kleinberg. Geographic routing using hyperbolic space. In *INFOCOM*, pages 1902–1909, 2007.

[11] S. Spiekermann, M. Rothensee, and M. Klafft. Street marketing: how proximity and context drive coupon redemption. *Journal of Consumer Marketing*, 28(4):280–289, 2011.

[12] P. K. Tat. A motivation-based model of coupon usage. *Journal of Promotion Management*, 3(1-2):1–14, 1996.