2013

Interest aware peoplerank: towards effective social-based opportunistic advertising

Soumaia Al Ayyat soumaia@aucegypt.edu

Interest Aware PeopleRank: Towards Effective Social-Based Opportunistic Advertising

Soumaia Al Ayyat
Computer Science and Engineering Dep.
The American University in Cairo
Cairo, Egypt
Email: soumaia@aucegypt.edu

Khaled A. Harras
School of Computer Science
Carnegie Mellon University Qatar
Doha, Qatar
Email: kharras@cs.cmu.edu

Sherif G. Aly
Computer Science and Engineering Dep.
The American University in Cairo
Cairo, Egypt
Email: sgamal@aucegypt.edu

***Abstract*—Various emerging context aware social-based applications and services assume constant non-disruptive connectivity. Mobile advertisers in such environments want to reach potentially interested users in a given proximity and within a specified short-duration, whether these users are connected to the network or not. While opportunistic forwarding algorithms can be leveraged for forwarding these advertisements, there is little incentive for those not interested in the ad to act as forwarders. Our goal in this paper is to leverage explicit interest, gathered from a user's social profile, and integrate it with social-based opportunistic forwarding algorithms in order to enable soft real-time opportunistic ad delivery in intermittently connected mobile networks. We propose IPeR, a fully distributed interest-aware forwarding algorithm that integrates with PeopleRank to reduce the overall cost and delay while reducing the number of contacted uninterested candidates. Our results, obtained via simulations and validated with real mobility traces coupled with user social data, are promising. In comparison to interest-oblivious socially-aware protocols such as PeopleRank, the IPeR approach reduces the cost to 70% to reach the same delivery ratio, and reduces the ratio of contacted uninterested forwarders by 23%. It also achieves an extra 70% recall and 107% accuracy with only 2% less precision.**

I. Introduction

The rapid advancements in mobile technologies and online social networks have paved the way for a new realm of context aware social-based applications and services. One of the most prominent service domains that have benefited from this is that of socially-aware advertising [1] [2] [3]. Most, if not all of the contributions in this area assume that targeted users are constantly connected to the network, and the types of advertisements shown are typically not time sensitive. This assumption is not always the case since edge wireless networks in crowded places like malls, train stations or theme parks may be costly to utilize, suffer from high medium contention, or simply unavailable at times. The goal for advertisers in such environments is to deliver ads to potentially interested users within a given proximity and within a specified short-duration, whether these users are connected to the network or not.

While opportunistic networking techniques is an attractive option that would enable advertisers to fulfill this goal, current solutions fall short in fulfilling this goal in different ways. Earlier solutions are generally focused on various forms of controlled flooding to reach a single destination within an intermittently connected network [4] [5] [6]. Other solutions have taken social information such as the users social rank, betweenness or centrality into account to further reduce cost without sacrificing delay in order to reach this single destination [7] [8] [9]. More recently, researchers have targeted group destinations that would match certain profiles [10] or common inferred interest amongst group members [11]. These solutions, however, either require high computation and storage capacities, assume the ability to maintain mobility pattern history or do not scale well with the increase in number of users. More importantly, these solutions assume the full cooperation of forwarding nodes in the network.

We take the first steps in this paper to leverage explicit interest, which can be easily gathered from a users social profile, and integrate it with social-based forwarding algorithms in order to enable soft real-time opportunistic ad delivery in mobile networks. We avoid engaging nodes that would not be interested in a given message as an incentive mechanism to participating users who would then mainly forward ads if they happen to be interested in its content. More specifically, we build upon a sample socially-aware forwarding algorithm, PeopleRank [7], by integrating interest obtained from online social profiles into the opportunistic forwarding algorithm. We consequently develop IPeR, a fully distributed interest aware social-based algorithm that leverages Jaccard set similarity between user interest vectors and advertisement specialization interest vectors. IPeR is essentially designed to magnify a nodes social rank if a given node and its friends are interested in a given message or advertisement. The consequence of our algorithm is to reach a subset of nodes that would be interested in a particular message, while utilizing mainly forwarder nodes that would be potentially interested in this message.

We evaluate IPeR via simulation in mall environments with realistic mobility patterns using the Self-similar Least Action Walk (SLAW) mobility model [12], and further validate our results using real social-based mobility traces gathered at the INFOCOM and SIGCOMM conferences [13] [14]. We measure the delivery ratio, cost, delay and effectiveness of our algorithm, and compare its performance to Epidemic Routing [4], used as bounding benchmark, PeopleRank [7], and a simple Interest-based forwarding algorithm we developed as a bound for interest-based algorithm performance. Our results show how accommodating interest in social-based opportunis-

978-1-4673-5939-9/13/$31.00 ©2013 IEEE

tic forwarding algorithms outperforms other interest unaware forwarding algorithms. For example, IPeR outperforms PeopleRank in cost by 70%, recall by 70%, and accuracy by 107% with only 2% less precision. More importantly, IPeR reduces the ratio of contacted uninterested forwarders by 23%.

II. Related Work

Most relevant to our work, are the various social-based forwarding algorithms that have been proposed in the opportunistic networking community targeted towards specific destinations or towards general profile matches. Algorithms targeting specific destinations include PeopleRank [7], HiBOp [15], BubbleRap [9] and SimBet [8]. These algorithms exploit user social 'betweenness' or centrality metrics to identify the most appropriate nodes for efficiently forwarding messages to a specific destination. These solutions, however, are targeted towards a given destination with a known address and produce high network overhead due to their periodic exchange of context information [16]. Since our solution is based on the PeopleRank [7] algorithm, we discuss it in details in the following section.

On the other hand, algorithms targeting groups based on profile matches rather than specific nodes include Socialcast [11] and Profilecast [10]. Socialcast infers interest commonality based on overlap in community memberships or mobility patterns, while Profilecast uses association matrices to identify commonality in user behavioral profiles that are generated based on the users mobility history. However, these solutions either require high computation and storage capacities, assume the ability to maintain mobility pattern history, or do not scale well with the increase in number of users or their interests.

Apart from the drawbacks mentioned above, our goal is for socially-aware forwarding algorithms that disseminate advertisements to mainly target interested users without taxing uninterested users by relying on them for this dissemination. None of the solutions mentioned above take into consideration the significance of the intermediate nodes' interest as an incentive to forward messages/advertisements. Furthermore, IPeR enables better privacy compared to previous work, as all the computation occurs in a distributed mode at each node's premises. Only friends can exchange social information and interest vectors, while others only exchange ad interest vectors and resulting values. These values cannot be used to infer the node's interest vector, friends list, or their encounters record.

III. Integrating Interest with PeopleRank

In this section we first provide a brief overview of the PeopleRank algorithm, and then describe our new interest-based algorithm, IPeR, along with the assumptions made in order to realize this system.

A. Overview of PeopleRank

PeopleRank is a recently introduced message forwarding algorithm based on forwarding messages utilizing the socially popular people nodes in place [7]. The algorithm is based on the hypothesis that socially popular nodes form better candidates to deliver messages to destinations given that there is a higher probability that such nodes will encounter destinations more quickly. Nodes are socially ranked as per their social relationship whether through a declared friendship, or if they share common interests.

To achieve a balance between social based forwarding and opportunistic forwarding, PeopleRank introduces a damping factor that decides on the percentage of reliance on the social ranking versus the opportunistic encounter of carriers. A damping factor (d) of value 0 in PeopleRank indicates total reliance on opportunistic forwarding, while a value 1 indicates total reliance on social based forwarding. Values between 0 and 1 indicate the weight given to each of the two approaches respectively. Empirical runs demonstrate that the optimal value for the damping factor d is 0.87 [7].

There is a contact-aware PeopleRank version (CA-PeR) that ranks nodes using their social rank and social activeness. Node activeness is measured by how frequent a node encounters social contacts. In effect, a node's social rank is either rewarded or penalized by the count of the node's encounters with its contacts as represented in the following equation [7].

$$CA-PeR(i) = (1-d)+d* \sum_{j\in F(i)} \frac{CA-PeR(j) * w_{i,j}}{|F(j)|} \tag{1}$$

where d is the damping factor, $CA-PeR(j)$ is friend j's PeopleRank value, $F(i)$ is the set of friends of node i who are available in vicinity, and $w_{i,j}$ is the contact-component.

$$w_{i,j} = \frac{|encounters_{i,j}|}{\sum_{k\in F(i)} |encounters_{i,k}|} \tag{2}$$

The CA-PeR value consists of a pure opportunistic component $(1-d)$, a contact-aware component w and a social-aware ranking component. The social-aware component is based on the concept that the node has a high social rank if on average its friends have high social ranks. This component is computed by a sum of the CA-PeR values of all the friends j of the node i averaged by the count of those friends $|F(i)|$.

B. Interest-aware PeopleRank

We explore the effect of integrating social interest with the forwarding process of the CA-PeR version. We introduce another parameter in ranking the nodes besides the typical social ranking and activeness used in the CA-PeR version. In specific, to consider a node for forwarding a message such as an advertisement, we compute the similarity in interest between the candidate forwarding node and the forwarded advertisement message, and use this information for further decision making. We add an interest similarity parameter to the CA-PeR equation to not only accommodate social ranking, but also an "interest-aware" social ranking component. The damping factor (d) used in CA-PeR will also still determine the amount of reliance on opportunistic forwarding and the new interest-aware social ranking component we coined. Interest-aware social ranking consists of the usual social ranking component of CA-PeR which will now be rewarded if the there is interest similarity in place, and penalized otherwise.

After eliciting the interest of nodes and of the advertisement message, and depicting it in the form of an interest vector, similarity is then measured using the Jaccard set similarity index [17]. If similarity is above a certain threshold, the rank of candidate nodes for forwarding is rewarded, and de-rewarded otherwise. The higher the rank of a node and its contacts, the more likely a node becomes a candidate for forwarding the message. A candidate node is highly ranked if it is linked to popular friends and also if the node and its friends are interested in the advertisement message. The following equation formalizes this concept.

$$IPeR(i) = (1-d) + d * \sum_{j \in F(i)} \frac{CA-PeR(j) * w_{i,j} * PSInt(j,Ad)}{|F(j)|} \quad (3)$$

where PSInt(j,Ad) is the Penalized Similar Interest

$$PSInt(j,Ad) = \begin{cases} (SInt(j,Ad) + reward) * 100 & \text{if} SInt \geq thr_{Int}, \\ (SInt(j,Ad) - penalty) * 100 & \text{otherwise} \end{cases} \quad (4)$$

where $SInt(j,Ad)$ is SimilarInterest and computed by the Jaccard set similarity between the user's interest vector and the ad specialization interest vector. Thus, IPeR magnifies the node's social rank if the node and its friends are interested in the message above a certain interest threshold thr_{Int} and it is penalized otherwise.

According to the logic of the IPeR Algorithm, first, the nodes that initiate an advertisement message (advertisers) rank the users in proximity using the IPeR function based on their candidacy to forward the advertisement message. The advertiser then sends the message to the 'interested forwarders' whose Similarity Interest $SInt(j,Ad)$ is beyond a certain threshold. As these forwarders encounter other nodes, they check their interest and social rank, whether they have already received this message or not, and their willingness to forward it to others along their way. In case of match, they forward the message to the new candidates. This process is repeated until the target time duration t expires or the target number of recipients is achieved.

Algorithm 1 Distributed IPeR Algorithm

Require: $|F(i)| \geq 0$, SInt(source,Ad) = 0
1: $IPeR(i) \leftarrow 1-d$
2: $\forall$ time t every n seconds
3: **while** i is in contact with j **do**
4:   **if** $j \in F(i)$ **then**
5:     send($IPeR(i), |F(i)|, FV_{interest}(i)$)
6:     receive($IPeR(j), |F(j)|, FV_{interest}(j)$)
7:     update($IPeR(i)$) (Eq. 3 and 4)
8:   **end if**
9:   **while** $\exists Ad \in$ buffer(i) **do**
10:     send($FV_{interest}(Ad)$)
11:     receive($SInt(j,Ad)$)
12:     **if** SInt(j,Ad) $\geq$ Destination-Interest-Threshold **or** (SInt(j,Ad) SInt(i,Ad) **and** IPeR(j) $\geq$ IPeR(i)) **then**
13:       Forward(Ad, j)
14:     **end if**
15:   **end while**
16: **end while**

Initially, all the nodes' IPeR values favor opportunistic forwarder selection i.e. IPeR = (1-d). Whenever two nodes come in contact, if they are friends, they exchange their IPeR values, the count of each one's friends $|F(i)|$, and their interest feature vectors $FV_{interest}(i)$ to update their IPeR ranks as per the equations 3 and 4 (lines 3-8). Whenever an ad holder i comes in contact with another node j, they exchange their current IPeR ranks, and node i sends the ad interest vector FV_{Ad} to node j to receive the computed $SInt(j,Ad)$ (lines 9-11). If node j belongs to the destination set of this ad (line 12), node i forwards to node j a copy of the ad. Also, if $IPeR(j) \geq IPeR(i)$ and $SInt(j,Ad) > SInt(i,Ad)$, node i forwards to node j a copy of the ad (lines 12-13).

C. System Realization Assumptions

For our algorithm to operate effectively, a set of attainable assumptions given today's technology are made. These assumptions include the presence of an ontology of interest among the nodes of interaction. We also assume that each user node has an installed client that carries a local copy of the user's social profile cached from his online social network. We assume direct interest is extracted from the social profile of the candidates, and that all the messages have the same size for simplicity of cost calculations.

This algorithm also considers short-duration advertisements that target users located in a place within a short period of time. Furthermore, the algorithm does not assume the existence of a fully connected social graph among the users in place, in contrast to CA-PeR algorithm, since it is based on interest and friendship and to be applicable in a mall environment within a mobility duration of one hour. We have simulated both algorithms once with a fully connected social graph and another time without this precondition to prove their applicability in either environment.

IPeR considers the message sender to be the source which has no interest in receiving its own message. Thus, its IPeR value starts as (1-d) and never improves due to the zero-interest component in the equation. However, all the forwarders update their IPeR values as they encounter their friends to become selective based on the candidate's popularity and interest.

IV. Evaluation

In this section we evaluate our proposed algorithm via simulation, and validate our results using real social-based mobility traces. We briefly describe our setup, and propose a subset of our results.

A. Simulation Setup and Parameters

We build our own simulator for a mall environment with 20 shops randomly distributed over an area of 1000m x 1000m. For accurate mobility patterns, we import user traces from the Self-similar Least Action Walk (SLAW) mobility model [12]. SLAW implements social contexts present among people sharing common interests in small scale communities such as university campuses, malls, or theme parks. To experiment with various conditions, we vary user density from 20 to 300, and vary their device ad hoc wireless range from 10m to 100m as shown in Table I. We run simulations for up to 12 hours to fully understand the system's behavior, but are only interested in the system's performance during the first hour, since our focus is on relatively short-time ads that target users during a single visit. Our results are based on ads generated by 2 of the 20 shops. All our results are shown as an average of 20 runs changing the random distribution of the users' mobility, profiles and friends list. Table I lists the most prominent parameters of our SLAW-based simulation environment.

In reality, not all users are interested in the same ads. We accommodate this factor by appropriately setting the similarity interest of a certain percentage of the users $SInt(interestednode, Ad) \geq 0.5$ with $SInt(destinationnode, Ad) \geq 0.9$. The remaining set of interested candidates are considered interested eligible carriers that would forward the ad to the destination nodes. While we test various user interest distributions, we only show results for the discrete uniform distribution; users are equally distributed between 11 categories with varying interest rates ranging from 0 to 1. Accordingly, the destination set constitutes 18% of the nodes while the interested forwarders cover 36%.

Our simulator generates random social profiles with random privacy settings per interest for each user. Furthermore, the constructed friendship social graph includes up to 20% of the available users in the friend list for each user. We also set the divergence between the encounters graph and the synthesized social graph as per the distance-based heuristic shown in previous work [7] to be around 0.2; this is an acceptable measure close to the optimum divergence value shown in that paper.

Finally, we have conducted experiments with various values of SInt(source,Ad) to deduce its effect on the algorithm performance as it acts as a starting cutoff point for forwarders' selection. The SInt(source,Ad) acts as a knob controlling the acceptable set of contacted uninterested nodes.

TABLE I
SIMULATION ENVIRONMENT PARAMETERS

Parameter	Nominal Value	Range
No. of users	50	20 - 300
No. of shops	2	1 - 20
Set of Interests	10	5, 10
Similarity interest distribution	discrete uniform	normal, discrete uniform
Destination set	18%	10% - 50%
damping factor	0.87	d = 1 - div [7]
SInt(source,Ad)	0.3	0 - 1

B. Validation with Real Traces

We validate the performance of our algorithm using real mobility traces gathered during the INFOCOM 2006 conference [13]. In that experiment, 20 static iMote nodes were installed to detect mobile devices within a 100m range, and 78 conference attendees were given iMotes with a 30m range. We import this INFOCOM06 dataset into our simulator and synthetically generate user profiles, friend lists, and interest feature vectors for the detected users. We use 2 of the static nodes to represent the shops generating ads, and show results within the same time frame of 1 hour.

For further validation, we import the mobility traces, interests and friendship graphs gathered during the SIGCOMM 2009 conference [14]. In this conference, 76 participants were handed in smartphones and were asked to use the installed MobiClique application for mobile social networking during the conference. Their social information, namely the list of friends and interests - was collected from their Facebook social profile. Thus, this dataset provides real friendship and interest graphs. We pick any of the users to be the source of the ads and show the results within 1-hour time frame. For space limitation, we present SLAW and SIGCOMM results only.

C. Simulation Metrics

The goal in our work is for advertisers to opportunistically reach the most relevant/interested users in the least amount of time possible. This goal should be attained while minimizing the overall cost, especially for users not interested in the ads. To evaluate the effectiveness of delivering on-time advertisement within a shopping mall, train station, or airports, we examine the above metrics in forwarding ads to interested users. We measure the performance of the compared algorithms to achieve this level of effective advertising using the following metrics:

Delivery Ratio: This is the percentage of nodes in the matching destination node set that were successfully reached. We also measure the delivery ratio as a percentage reached from the *total* node set in order to evaluate the algorithms under varying destination interest thresholds.

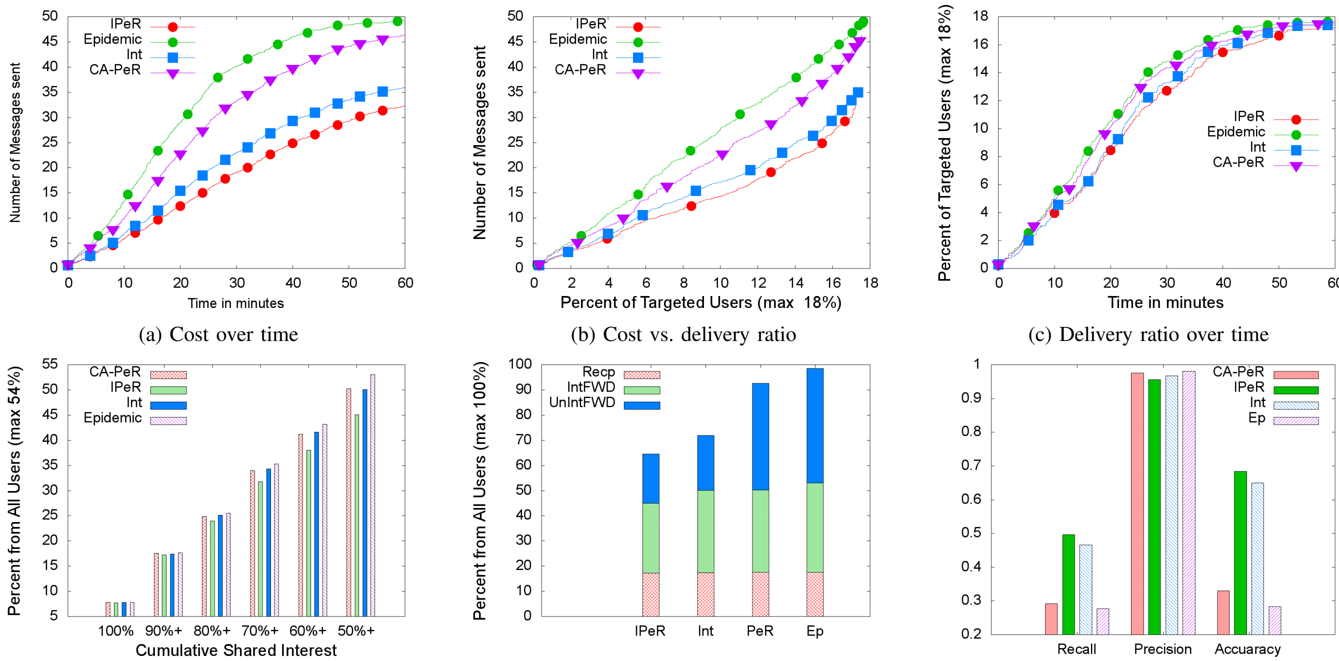
Cost: This is measured by the total number of message replicas that have been generated at any given time. We also measure the cost per unit delivery ratio as a tangible reflection of algorithm efficiency.

Delay: Another reflection of efficiency is that instead of showing the overall delay in time, we indicate the percentage of contacted destinations at various time stamps. In other words, we show the amount of time taken to reach the least common delivery ratios.

Effectiveness: In our domain, an algorithm is effective if it contacts a high portion of the interested nodes while simultaneously avoiding the uninterested ones. We measure this by the ratio of contacted nodes classified by their interest; nodes are either interested forwarders, destination nodes, or uninterested forwarders. We also measure this effectiveness through recall, precision, and accuracy [18].

D. Results

In this section, we examine the effectiveness of our interest social-based algorithm, IPeR, by comparing its performance to



(a) Cost over time

(b) Cost vs. delivery ratio

(c) Delivery ratio over time

(d) Delivery ratio vs. interest threshold

(e) Effectiveness: Interest-based node classification

(f) Effectiveness: Recall/Precision/Accuracy

Fig. 1 SLAW Mobility - SInt(source,Ad) = 0

a sample social-contact aware algorithm, CA-PeR, an Interest-Only algorithm, and a social-unaware Epidemic algorithm that is used mainly for performance bounds. We note that the Interest-Only algorithm relies on the similarity interest metric for forwarder/destination node selection without consideration for social links and activeness. We only share a representative set of results due to space limitation. The SLAW-based and SIGCOMM09-based experiment results are illustrated in Figures 1 and 2 respectively.

It is worth noting that the increase in SInt(source,Ad) value significantly decreases the set of contacted uninterested nodes. For space limitation, we demonstrate IPeR's base performance when SInt(source,Ad) = 0 with SLAW data, and present the variation in performance when SInt(source,Ad) = 0.3 in SIGCOMM results.

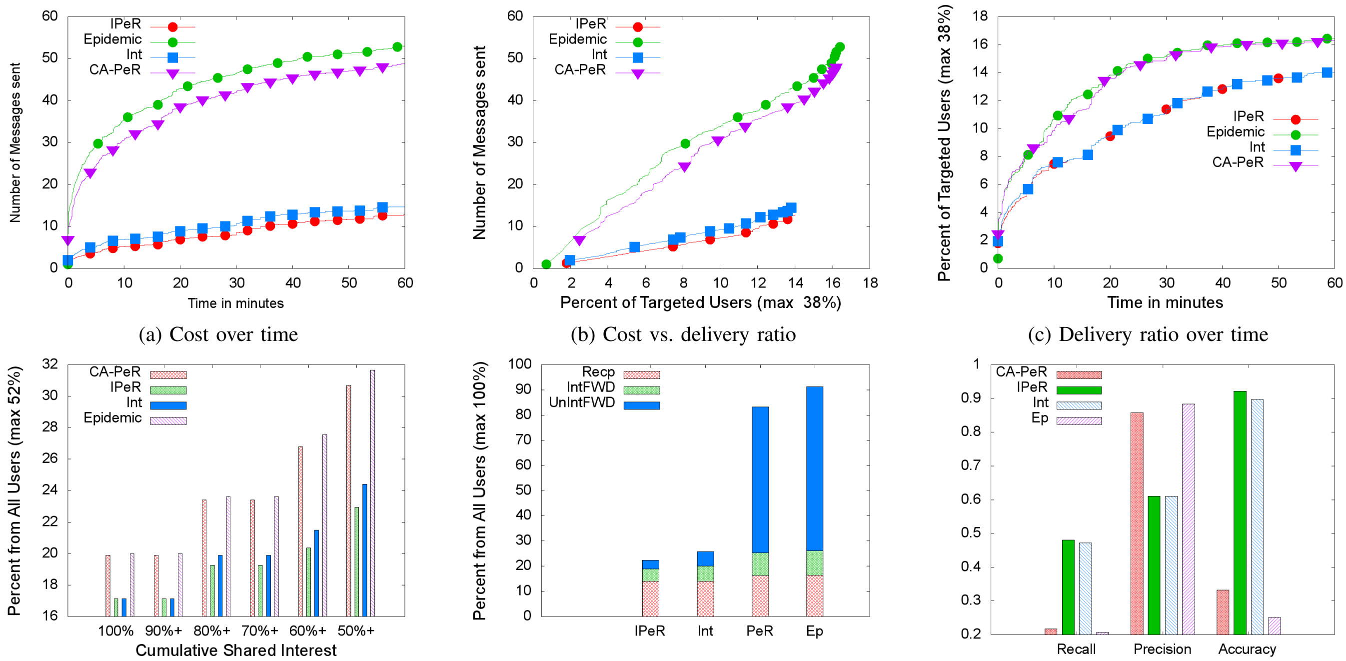
The cost time analysis of the four algorithms delivering the message to the destination set, consisting of 18% of all the users, is illustrated in Figures 1(a) and 2(a); the cost to delivery ratio is also shown in Figures 1(b) and 2(b). Overall, we observe that IPeR reduces the cost to 70% of that incurred by CA-PeR while reaching all nodes in the destination set CA-PeR contacted within 50 minutes. Thus, IPeR only costs 1.87 messages per unit delivery ratio, mainly due to the fact that IPeR avoids forwarding messages to uninterested forwarders. We also highlight the tradeoff between the reduction in cost and the extra delay to reach the target set; based on Figures 1(c) and 2(c), IPeR and Interest-Only consume an extra 14.4% delivery time when compared to CA-PeR.

The success delivery ratio achieved by each algorithm at several cumulative destination set similarity interests is represented by Figures 1(d) and 2(d). The results indicate that Interest-Only succeeds to reach a slightly higher percentage within the destination sets compared to IPeR. Both algorithms reach a slightly lower percent compared to CA-PeR despite the restriction imposed on them by interest-based forwarding.

The crucial metric of measuring the percentages of the contacted interested forwarders, uninterested forwarders, and destination nodes is illustrated in Figures 1(e) and 2(e). IPeR and Interest-Only significantly reduce the percent of contacted uninterested forwarders by 23% compared to CA-PeR. In addition, they succeed to contact almost the same ratio of interested forwarders and destination nodes as those contacted by Epidemic and CA-PeR. We note that, the Interest-Only algorithm succeeds to contact a slightly higher percentage of interested forwarders compared to IPeR since it concentrates only on interest in selection.

The final metric related to recall, precision, and accuracy is illustrated in Figures 1(f) and 2(f). Although IPeR and Interest-Only provide less recall ratio in comparison to CA-PeR and Epidemic, they achieve 70% extra recall and 107% extra accuracy with comparable precision. It is worth noting that both algorithms lose 28.8% precision in comparison to CA-PeR in the SIGCOMM results due to forwarder selection restriction upon the increase in SInt(source,Ad).

Overall, our results show that IPeR significantly reduces cost mainly by avoiding the majority of the uninterested candidates. It also maintains the same success delivery ratio, and outperforms in recall and accuracy by applying the same CA-PeR ranking approach in addition to setting a reward/penalty score to this rank. While it seems like Interest-Only has a comparable performance compared to IPeR, other experiments conducted by us have shown that IPeR does largely better if the interest distribution is not as smooth as the chosen uniform distribution. The reason for this result is that IPeR imposes a



(a) Cost over time (b) Cost vs. delivery ratio (c) Delivery ratio over time

(d) Delivery ratio vs. interest threshold (e) Effectiveness: Interest-based node classification (f) Effectiveness: Recall/Precision/Accuracy

Fig. 2 SIGCOMM09 Results - SInt(source,Ad) = 0.3

balance between interest and social information, and hence, any discrepancy in interest information availability severely impacts the performance of Interest-Only compared to IPeR.

V. Conclusion and Future Work

In this paper, we have taken the first steps towards showing the impact of incorporating interest in socially-aware opportunistic forwarding algorithms, in order to accommodate soft real-time advertisement dissemination to disconnected mobile users. We have introduced the IPeR algorithm, which built interest awareness into the PeopleRank, a representative socially-aware forwarding algorithm. Overall, our evaluation via simulation and real data trace based experiments demonstrate the promising gain in efficiency and effectiveness of interest-aware social-based algorithms, especially in non-uniformly distributed interest communities.

For future work, we intend to construct a generalized framework based on this work that incorporates interest in any social-based forwarding algorithm. We also need to incorporate better incentive mechanisms that would further encourage users to forward messages. Finally, we plan to better evaluate the performance of this framework on large-scale real data that mixes mobility patterns with real social profiles and interests.

References

[1] H. Haddadi, P. Hui, T. Henderson, and I. Brown, "Targeted Advertising on the Handset : Privacy and Security Challenges," in *Pervasive Advertising*. Springer London, 2011, pp. 1–19.

[2] E. D. Spiegler and F. Michahelles, "Field Experiment of Social Network Content on Public Displays in Small-Space Retail Shops," in *3rd Pervasive Advertising Workshop*, 2010.

[3] E. D. Spiegler, C. Hildebrand, and F. Michahelles, "Social Networks in Pervasive Advertising and Shopping," in *Pervasive Advertising*. Springer London, 2011, pp. 207–225.

[4] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," CS-200006, Duke University, Tech. Rep., 2000.

[5] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: An efficient routing scheme for intermittently connected mobile networks," in *SIGCOMM05 Workshops*, 2005.

[6] A. Lindgren *et al.*, "Probabilistic Routing in Intermittently Connected Networks," in *Service Assurance with Partial and Intermittent Resources*. Springer Berlin, 2004, vol. 3126, pp. 239–254.

[7] A. Mtibaa, M. May, and M. Ammar, "Social Forwarding in Mobile Opportunistic Networks: A Case of PeopleRank," in *Communication and Social Networks*. Springer, 2012, vol. 58, pp. 387–425.

[8] E. M. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs," in *8th ACM MobiHoc Proceedings*, 2007, pp. 32–40.

[9] P. Hui *et al.*, "Bubble Rap: Social-based forwarding in Delay Tolerant Networks," in *9th ACM MobiHoc Proceedings*, 2008, pp. 241–250.

[10] W. jen Hsu, D. Dutta, and A. Helmy, "Profile-cast: Behavior-aware mobile networking," in *IEEE WCNC*, 2008, pp. 3033 –3038.

[11] P. Costa *et al.*, "Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 748–760, Jun. 2008.

[12] I. Rhee *et al.*, "CRAWDAD data set ncsu/mobilitymodels (v. 2009-07-23)," Downloaded from http://crawdad.cs.dartmouth.edu/ncsu/mobilitymodels.

[13] J. Scott *et al.*, "CRAWDAD trace cambridge/haggle/imote/infocom2006 (v. 2009-05-29)," Downloaded from http://crawdad.cs.dartmouth.edu/cambridge/haggle/imote/infocom2006.

[14] A.-K. Pietilainen and C. Diot, "CRAWDAD data set thlab/sigcomm2009 (v. 2012-07-15)," Downloaded from http://crawdad.cs.dartmouth.edu/thlab/sigcomm2009.

[15] C. Boldrini, M. Conti, and A. Passarella, "Exploiting users social relations to forward data in opportunistic networks : The HiBOp solution," *Pervasive and Mobile Computing*, vol. 4, pp. 633–657, 2008.

[16] H. A. Nguyen and S. Giordano, "Context information prediction for social-based routing in opportunistic networks," *Ad Hoc Networks*, vol. 10, no. 8, pp. 1557 – 1569, 2012.

[17] R. Vernica, M. J. Carey, and C. Li, "Efficient parallel set-similarity joins using mapreduce," in *ACM SIGMOD Proceedings*, 2010, pp. 495–506.

[18] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness & correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.