Extending the Reach of Social-Based Context-Aware Ubiquitous Systems

Dania Abed Rabbou, Abderrahmen Mtibaa, Khaled A. Harras
School of Computer Science
Carnegie Mellon University
{dabedrab,amtibaa,kharras}@cmu.edu

I. Introduction

Social networking has recently gained unprecedented popularity among Internet users. The advent of smart phones equipped with mobile technology such as WiFi, 3G, 4G, and sensors, has exponentially increased the number of mobile phone users thus enriching the content of online social networks such as Facebook. This has increased the importance of social- and context-aware systems that leverage social information with location awareness to provide personalized services. Figure 1 illustrates how such systems renders Bobs shopping activity in a mall an exciting endeavor. Bob receives notifications based on his Facebook profile and location in positions A, B, and C. He is able to tune the world around him to his liking, discover surrounding people and places, and seize various social/business opportunities.

To receive context-aware recommendations in real-time, mobile devices must be connected to wireless networks. However, in reality, many devices are not connected all the time due to the absence of coverage or high energy consumption and cost (like 3G service). Moreover, wireless connectivity is intermittent in heterogeneous environments in which users move from one network to another and can suffer from contention under high load.

Previous systems in this domain were based on centralized and distributed paradigms. Serendipity [8], Live Social Semantics [1], SPETA [9], and SocialFusion [4] are centralized systems that collect and store users social information in a server to provide context-aware recommendations about other users within Bluetooth range. While a centralized paradigm guarantees timeliness, it assumes persistent connectivity of users at all times. As an alternative paradigm to achieving the same functionality, WhozThat [3], Ad-Hoc Smart Spaces [5] and MobiClique [12] proposed a Bluetooth-based distributed architecture in which devices obtain context-aware messages opportunistically using intermediate nodes. Successfully overcoming the connectivity assumption, such a paradigm compromises on the extent of information dissemination and timeliness because delivery is solely dictated by user mobility and meetings.

The shortcomings of existing paradigms call for a new hybrid paradigm that enables the reachability to unconnected users while still maintaining timeliness. Therefore, among the various critical challenges in the domain of social-context-awareness as depicted in Figure 2, we choose to address the end-to-end communication challenge.

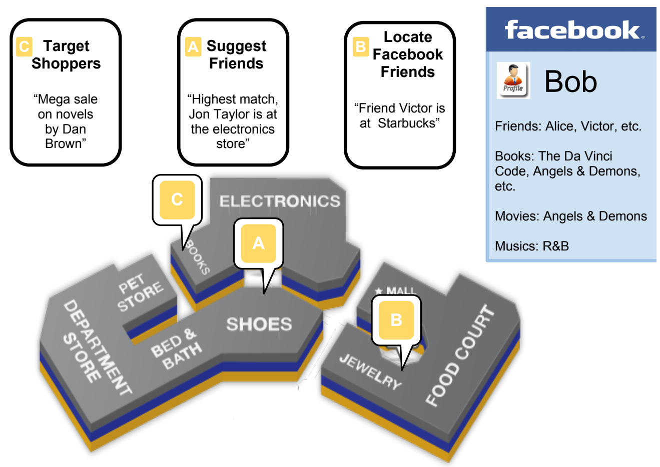


Fig. 1. Bob is shopping in a mall and receiving real-time context-aware recommendations based on his on-line social-network profile.

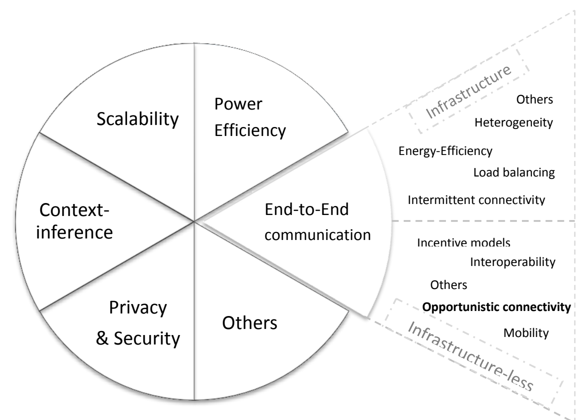
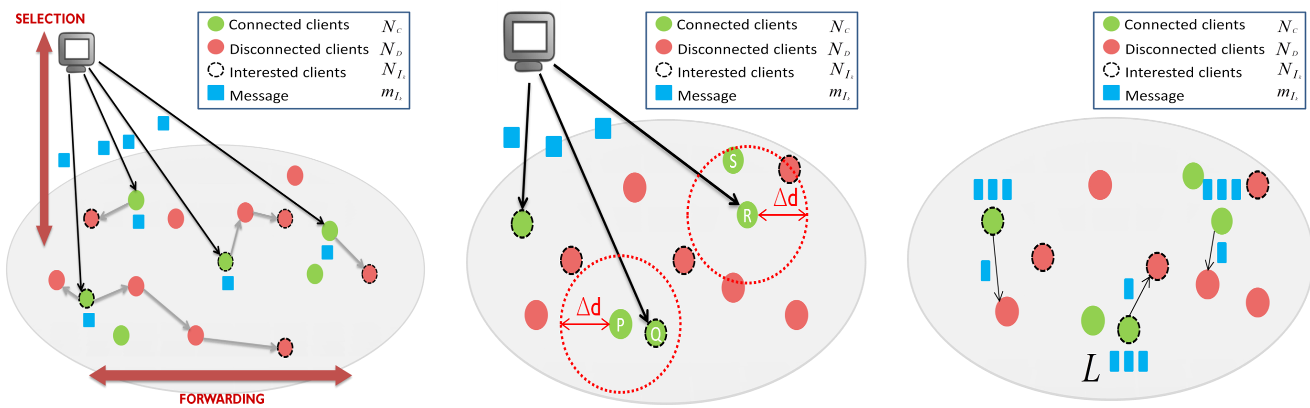


Fig. 2. Challenges in the domain of social context-awareness and the challenge we address, namely *end-to-end communication*. Communication either assumes and requires a fixed infrastructure or is infrastructure-less. Each has a set of drawbacks and therefore we propose a hybrid paradigm.

II. Problem Definition

End-to-end communication generally refers to the transfer of data packets from a source to a destination. In the domain of ubiquitous systems, the packets are context-aware messages



(a) Selection and forwarding (b) Selection by the server (c) Forwarding by the message-carrying nodes

Fig. 3. Illustration of the selection and forwarding phases

that originate from a server or devices and are destined to devices whose users are interested in the message. The challenge is to deliver the messages to the maximum number of connected and unconnected users in the shortest possible time and minimum cost.

More formally, Let N be the set of all users registered with a central system S. N is partitioned into two sets N_C, the set of connected users, and N_D, the set of unconnected users. *Given an interest I_k and a related message m_{I_k} generated by S at time t, we need to deliver the message to $N_{I_k} \subseteq N_C \cup N_D$, the set of users interested in I_k, with minimum delay δt and minimum number of message replicas.*

Much research has been done for minimizing the delay in end-to-end communication between a user and a server in an infrastructure-based environment. Therefore, connected users can directly receive messages with the minimal delay.

Unconnected users, however, can acquire messages opportunistically using an infrastructure-free paradigm. *Opportunistic communication* refers to the exchange of messages between users when they come in communication range of one another. The research community has proposed general techniques for opportunistic communication which we discuss in the following section. Note that in the following sections, we refer to users as nodes.

III. Related Work

Opportunistic communication techniques can be classified into two categories: Non-social and Social.

A pioneer work in the non-social category is Epidemic Routing [14] in which a message-carrying node transfers a replica of the message to every node it encounters within communication range. Essentially, this technique unconditionally floods a message throughout the entire network. While, it overflows node buffers with messages, it achieves the minimal delay to reach destination nodes. Evolutions of Epidemic Routing are MaxProp [6] which prioritizes messages to flood and Spray and Wait [13] which sprays or transfers only some L replicas of a message to nodes within communication range and waits until one of the nodes meets the destination. In order to reduce the cost or number of replicas, RAPID [2] and PRoPHET [16] proposed probabilistic techniques in which nodes transfer replicas only to other nodes that have a high probability of encountering a destination. This information is obtained from the history of encounters. Another technique is Message Ferry [15] that deploys ferries (like robots or vehicles) to carry and deliver messages between nodes.

The category of social techniques is a recent one which exploits social-relationships among nodes to identify to transfers replicas to most popular nodes. In SimBet [7] and BubbleRap [10], each node is assigned a centrality value which is equal to the number of its social friends and message replicas are transferred to nodes with high centrality. PeopleRank [11] is another social technique inspired by PageRank that gives weights to nodes according to the number of important nodes, nodes with high weights, they are socially linked to.

IV. Our Solution

To disseminate a context-aware message originating from a server to a set of interested connected or unconnected nodes, our solution consists of two phases: *Selection* and *Opportunistic Forwarding*. Opportunistic forwarding refers to the exchange of messages between nodes when they come in communication range. While selection is performed by the server, opportunistic forwarding is performed by all the message-carrying nodes as shown Figure 3 (a).

A. Selection

We assume that the server, S, stores the social profiles consisting of basic information, friends and interests of all registered nodes N. S also builds and maintains a social graph of N nodes in which a direct link exists between two nodes if they are listed as friends in their social profiles. Using this graph, S computes the betweenness centrality as:

5 4 3 2 1	Technique	Number of replicas	End-to-End delay	Success rate	Assumptions	Forwarding Complexity	Score (25)
Epidemic	Flooding	V. High	Optimal	Optimal	Unlimited Bandwidth & buffers	None	11
MaxProp	Prioritized Flooding	V. High	Low	High	Unlimited Bandwidth	None	13
Spray & Wait	Limited Flooding	V. Low	High	V. High	Number of nodes	None	16
RAPID	Probabilistic: expected delay	High	Low	High	Inter-node meeting	High	11
PRoPHET	Probablistic: predict future meeting	High	Low	High	Unlimited Bandwidth	High	12
Message Ferry	Assisted Mobility	V. Low	High	V. High	Mobility	None	15
SimBet	Social Centrality	High	V. Low	V. High	Social neighbors	High	16
Bubble Rap	Social Centrality & Clustering	High	V. Low	V. High	Social Graph	High	14
PeopleRank	PeopleRank	High	V. Low	V. High	Social neighbors	Moderate	17

Fig. 4. Qualitative comparison between opportunistic techniques

$$\forall n \in N, C(n) = \sum_{i=1}^{N} \sum_{j=1}^{i-1} \frac{p_{ij}(n)}{p_{ij}} \quad (1)$$

where p_{ij} is the number of shortest paths between i and j and $p_{ij}(n)$ is the number of such paths that pass through n. Intuitively, the betweenness centrality measures the extent to which a node can facilitate communication between other nodes. At any time t, $\forall n \in N, n \in N_C \vee n \in N_D$. When a message m_{I_k} is generated by S, it computes the subset of interested nodes N_{I_k} as follows:

$$\forall n \in N, n \in N_{I_k} \quad if \; I_k \in Interests(n) \quad (2)$$

S sends the message to all the connected nodes in $N_{I_k} \cap N_C$. However, unconnected nodes in $N_{I_k} \cap N_D$ must obtain the message opportunistically from other message-carrying nodes. To increase the chances of message delivery, S transfers the message to additional connected nodes as follows (depicted in Figure 3 (b)). $\forall n \in N_C - N_{I_k}$:

- Find all nodes in the circular area whose center is n and radius is $\triangle d$.
- For each such node m,
  1) if m received m_{I_k} then a copy of the message exists in that area and hence n will not receive it. For example, node P does not receive the message because the interested and connected node Q had received it.
  2) if m did not receive m_{I_k} then n will receive a copy if $C(n) > C(m)$, otherwise m will receive it. For example, node R receives the message because $C(R) > C(S)$.

B. Opportunistic Forwarding

Once the selected nodes receive the message, they begin forwarding the message opportunistically so that it eventually reaches the interested nodes. We assume that all nodes have unlimited buffers and they are willing to be message carriers. We use one or more of the opportunistic communication techniques discussed in Section III. We qualitatively evaluate and compare the techniques based on five metrics as shown in Figure 4. Each metric is assigned a color according to its performance where yellow and red signify the optimal and worst performance and are assigned 5 and 1 respectively. The overall performance of the technique is then computed as the sum of the values assigned to each metric.

We note that the social techniques, particularly PeopleRank, perform well in terms of end-to-end delay and success rate. However, the cost in terms of number of replicas can be improved. For this, we consider non-social techniques Spray-and-Wait and Message Ferry which have a very low cost. Therefore, we adopt the highest score non-social technique, Spray-and-Wait, and social technique, PeopleRank. We will combine these two techniques when performing forwarding as explained below.

Each node in the network computes its PeopleRank as follows:

$$PeR(n) = (1-d) + d \sum_{m \in SN(n)} \frac{PeR(m)}{|SN(m)|} \quad (3)$$

where $SN(n)$ is the set of nodes that node n physically encountered, referred to as social neighbours, and d is a damping factor defined as the probability that the social relation between nodes helps improve the rank of these nodes.

When message-carrying node n encounters another node m i.e. they become within communication range, the following steps occur as shown in Figure 3 (c):

- m sends its list of interests to n. If $I_k \in Interests(m)$, n sends a replica to m. Otherwise, m is not interested in the message but can act as a carrier.
- m sends the list of messages in its buffer and its PeopleRank. n sends a replica to m if $m_{I_k} \notin Messages(m)$ and $PeR(m) > PeR(n)$. A higher PeopleRank value indicates that a node is more socially linked and is hence more likely to meet the destinations.

All message-carrying nodes repeat this process of forwarding until they have each sprayed a total of L copies of the message in the network.

V. Evaluation

In this section we explain our evaluation set-up, metrics and parameters to evaluation our proposed solution.

A. Set-up

For evaluating our solution we built a centralized social-based context-aware system, called SCOUT, that provides real-time recommendations or messages to its registered users based on their social profile and current location. he system is based on a client-server architecture where the server, shown in Figure 5, consists of a task manager, a recommendation engine, and three databases that store social profiles, location information, and client settings.

A new SCOUT client or user connects to Facebook, fetches its social profile including basic information, friends and

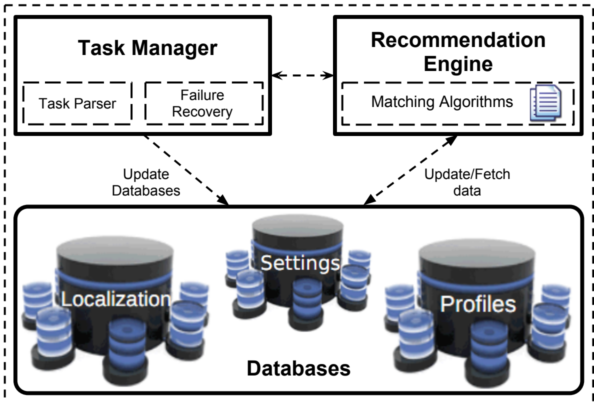


Fig. 5. High level architecture of SCOUT

interests, and sends it to the server. The multi-threaded server receives this information and delegates the task to the task parser. The parser parses the information and stores the data in the Profiles databases. As long as the client is connected, SCOUT communicates with a indoor localization engine to obtain the (x, y, z) coordinates of the client which it also stores in the location database. Besides the interests listed in the social profile, a client can set additional interests that are stored in the settings database.

The recommendation engine generates m_{I_k} for each interest k at rate α messages per unit minute. It then computes N_{I_k} and implements the selection algorithm as discussed in Section IV-A.

We then use a simulator to emulate the forwarding process as discussed in Section IV-B. We will use synthetic mobility traces generated by the simulator and real ones obtained from the online repository CRAWDAD [1].

B. Metrics and Parameters

We will measure three important metrics in the domain of social- and context-awareness namely **(i)** Cost: Total number of message replicas required to reach the destinations, **(ii)** End-to-end delay: Average time delay incurred to reach the destinations, and **(iii)** Success rate: Ratio of the number of destinations reached to the total number of destinations.

For each metric, we will vary the parameters **(i)** N_C: Number of connected nodes, **(ii)** $\triangle d$: Radius, **(iii)** L: Number of replicas to spray, and **(iv)** δt: deadline after which the message is dropped from the buffers.

References

[1] H. Alini, M. Szomszor, C. Cattuto, W. V. den Broeck, G. Correndo, and A. Barrat. Live social semantics. 2009.

[2] A. Balasubramanian, B. Levine, and A. Venkataramani. Dtn routing as a resource allocation problem. *SIGCOMM Comput. Commun. Rev.*, 37(4):373–384, 2007.

[3] A. Beach, M. Gartrell, S. Akkala, J. Elston, J. Kelley, K. Nishimoto, B. Ray, S. Razgulin, K. Sundaresan, B. Surendar, M. Terada, and R. Han. Whozthat? evolving an ecosystem for context-aware mobile social networks. *IEEE Network*, 22(4):50–55, 2008.

[1]http://crawdad.cs.dartmouth.edu/

[4] A. Beach, M. Gartrell, X. Xing, R. Han, Q. Lv, S. Mishra, and K. Seada. Fusing mobile, sensor, and social data to fully enable context-aware computing. In *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications*, HotMobile '10, pages 60–65, New York, NY, USA, 2010. ACM.

[5] A. Brodt and S. Sathish. Together we are strong— towards ad-hoc smart spaces. In *Proceedings of the 2009 IEEE International Conference on Pervasive Computing and Communications*, pages 1–4, Washington, DC, USA, 2009. IEEE Computer Society.

[6] J. Burgess, B. Gallagher, D. Jensen, and B. Levine. Maxprop: Routing for vehicle-based disruption-tolerant networking. In *Proceedings of IEEE Infocom*, 2006.

[7] E. M. Daly and M. Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '07, pages 32–40, New York, NY, USA, 2007. ACM.

[8] A. Foster and N. Ford. Serendipity and information seeking: an empirical study. *Journal of Documentation*, 59(3):321–340, May 2003.

[9] A. Garca-Crespo, J. Chamizo, I. Rivera, M. Mencke, R. Colomo-Palacios, and J. M. Gmez-Berbs. Speta: Social pervasive e-tourism advisor. *Telematics and Informatics*, 26(3):306–315, 2009.

[10] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: social-based forwarding in delay tolerant networks. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '08, pages 241–250, New York, NY, USA, 2008. ACM.

[11] A. Mtibaa, M. May, C. Diot, and M. Ammar. Peoplerank: social opportunistic forwarding. In *Proceedings of the 29th conference on Information communications*, INFOCOM'10, pages 111–115, Piscataway, NJ, USA, 2010. IEEE Press.

[12] A. K. Pietilinen, E. Oliver, J. LeBrun, G. Varghese, and C. Diot. Mobiclique: middleware for mobile social networking. pages 49–54, 2009.

[13] T. Spyropoulos, K. Psounis, and C. S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, WDTN '05, pages 252–259, New York, NY, USA, 2005. ACM.

[14] A. Vahdat and D. Becker. Epidemic routing for partially-connected ad hoc networks. Technical Report CS-2000-06, UCSD, 2000.

[15] W. Zhao, M. Ammar, and E. Zegura. A message ferrying approach for data delivery in sparse mobile ad hoc networks. In *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '04, pages 187–198, New York, NY, USA, 2004. ACM.

[16] Z. Zhou and L. M. Ni. Prophet address allocation for large scale manets. Technical Report MSU-CSE-02-20, Department of Computer Science, Michigan State University, East Lansing, Michigan, July 2002.