SimBetAge: Utilizing Temporal Changes in Social Networks for Pocket Switched Networks

Jó Ágila Bitsch Link, Nicolai Viol, André Goliath, Klaus Wehrle*
Distributed Systems Group
RWTH Aachen University, Germany
(bitsch|viol|goliath|wehrle)@cs.rwth-aachen.de

ABSTRACT

In this paper, we present SimBetAge, a delay and disruption tolerant routing protocol for highly dynamic socially structured mobile networks. We exploit the lightweight and ego-centric scheme of SimBet routing while at the same time taking the strength and the gradual aging of social relations into account and thereby increase the performance by one order of magnitude, especially in evolving network structures. We explore the model of similarity and betweenness over weighted graphs, and present a simulation on realistic traces from previous experiments, comparing our approach to the original SimBet, Epidemic Routing and Prophet.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Store and forward networks

General Terms

Algorithms, Design, Performance

Keywords

Routing, DTN, ego networks

1. INTRODUCTION

Pocket Switched Networks [9] or more generally Disruption/Delay Tolerant Networks are a representative of a new type of communication paradigm, in which data is delivered over eventually transportable dynamic networks. Data bundles are cached for a possibly long time before a suitable contact to a forwarding node is made. The proliferation of mobile devices leads us to the belief, that this paradigm will gain significant popularity in the future. The routing of data in this kind of networks can be categorized in several classes [5]. In the simplest case, a data bundle is epidemically forwarded in the network. More refined approaches forward either dependent on detailed plans of future contacts or some heuristics to predict the probabilities of those.

The identification of useful heuristics is therefore the core contribution of the majority of publications about routing in this kind of networks. The social structure has long been considered a good starting point for building heuristics in the community. However, the change over time of this structure, especially over long runs of time, has often been neglected. The contribution of this work is therefore the adaption of a previous social network based approach, namely SimBet [1], to consider the *progression* of social networks over time. To that end, the original definitions of *similarity* and *betweenness* had to be significantly adapted to deal with weighted graphs instead of binary ones. Additionally, we described a novel metric, the *directed ego flow betweenness*.

The remainder of this paper is structured as follows: After a short discussion on related work, we discuss SimBet in its original form and identify several of its short comings. Then, we present SimBetAge and discuss its structure and features in depth. We conclude with a thorough evaluation, showing that our approach performs up to 90% better than the original SimBet.

2. BACKGROUND AND RELATED WORK

One of the simplest approaches to use aging in MANETs is FRESH [2]. Interpreting the time of the last contact to a destination node as a measure of the distance, FRESH is able to discover short path in an efficient way. Although this approach needs virtually every node to encounter the destination at least once, it shows the importance of the age of a contact as a parameter for efficient routing.

In [12], Lindgren et al. presented a probabilistic routing scheme for intermittently connected networks called *PRoPHET* where each node keeps a table with probabilities for reaching every possible destination in the network. These probabilities depend on the frequency of observations, the time of absence and the distance to a destination, but does not make use of the social nature of the network.

Bubble Rap [10] explicitly discovers structural features by processing historical contacts and computing node communities and hierarchies a priory. However, Bubble Rap requires a setup period and does not adapt to dynamics in the network very well.

Our work builds upon previous work done by Daly et al.[1] who were one of the first to introduce social network analysis

*This research was funded in part by the DFG Cluster of Excellence on Ultra-high Speed Information and Communication (UMIC), German Research Foundation grant DFG EXC 89.

http://doi.acm.org/10.1145/1659029.1659034
U-NET'09, December 1, 2009, Rome, Italy.

in the context of delay tolerant networks, called *SimBet*. In this approach, routing decisions are based on two common properties of social networks, *social similarity* and *centrality betweenness*.

In contrast to PRoPHET, SimBet relies only on an ego-centric network view and is able to make efficient forwarding decisions even if the node has not encountered a path to the destination before. Their handling of node distances inspired us with respect to modeling the dynamics in relations as presented in section 3.1.

The calculation of ego betweenness in graphs with weighted edges was introduced by Freeman et al. [7]. We describe the approach in more detail in section 3.3. An extension for directed ego flow betweenness can be found in section 3.5.

2.1 Background on SimBet Routing

The motivation on SimBet is to use social network analysis to provide an efficient routing scheme that has no need for global knowledge about the network. Conceptually, a node forwards messages to a node that has a higher social similarity to the destination, i.e. it is more likely to have a contact to the destination. If there is no such node known to the forwarder, it instead forwards the message to nodes with a higher betweenness, i.e. to a potentially better forwarder to other parts/clusters in the network.

While similarity is defined on an ego-centric view, betweenness is defined on geodesics, the shortest paths between two points. Therefore, the calculation needs global knowledge. In [4] however the authors showed that the calculation of an ego-centric betweenness is still meaningful enough to fit a global representation.

The social structure can therefore be used as the macro structure for forwarding decisions. The similarity contributes strongest when a message is already near to the destination with respect to the social structure, whereas the betweenness has stronger effects when the forwarding node is far from the destination.

Accordingly, the similarity of u to v can be defined as the number of common encounters between u and v:

$$S_u(v) := |N_1(u) \cap N_1(v)| \tag{1}$$

The definition of the global betweenness centrality of u is therefore:

$$C_B(u) := \sum_{\substack{v \neq w \neq u \\ v,w \in \text{Nodes}}} \frac{g_{v,w}(u)}{g_{v,w}} \tag{2}$$

Thus, Daly defines the ego betweenness centrality by taking nodes v and w only from the neighborhood of u.

$$\theta_u := \sum_{\substack{v \neq w \neq u \\ v,w \in N_1(u)}} \frac{g_{v,w}(u)}{g_{v,w}} \tag{3}$$

2.2 Shortcomings on SimBet

Despite the fact that SimBet performs quite well as shown by the evaluations done by its authors [1], some weaknesses can be pointed out.Routing decisions in SimBet are based on comparing the social relations and roles of individual nodes at a specific time in a binary fashion. However, some relations are stronger than others, i.e. contacts between them are more frequent. Additionally, those relations may change over time. Therefore, a simple statement that x knows y without additional qualifiers is not enough.

$G(t)$	A time dependent graph
$u, v, w \in V$	The vertices u, v, w
$N_i(u)$	The i hop neighborhood of u
$e \in E$	An edge e
$\omega(e,t)$	The time-dependent weight or *freshness* of edge e, also abbreviated ω
$\omega(P,t)$	The freshness of a path P at time t
$\omega(u,v)$	The current weight between nodes u and v
α	A renewing/growth factor
γ	A aging/decay factor
Δt	Time steps since a previous contact
$g_{v,w}$	The number of geodesics between v and w
$g_{v,w}(u)$	The number of geodesics between v and w passing through u
$S_u(v)$	The similarity between u and v
$C_B(u)$	The betweenness centrality of u
$\theta(u)$	The ego betweenness centrality of u
$\sigma(u,v)$	The aged similarity between u and v
ϑ_u	The ego flow betweenness centrality of u
$\delta_u(d)$	The destination ego flow betweenness centrality of u with respect to the destination d
$U_v(d)$	The utility of v for transporting a message to d

Table 1: List of symbols and abbreviations used in this paper

Further, a node might have different roles: It may participate in different clusters (different circles of friends) so one node might be important for forwarding information from one cluster to another but not necessarily that important for forwarding within a cluster.

In figure 1 one can see how the network structure and the betweenness of individual nodes varies over time. Comparing its three subfigures it becomes evident, that the original model based on the ego betweenness centrality in a network is not adequate and will lead to inaccurate decisions for later transmissions.

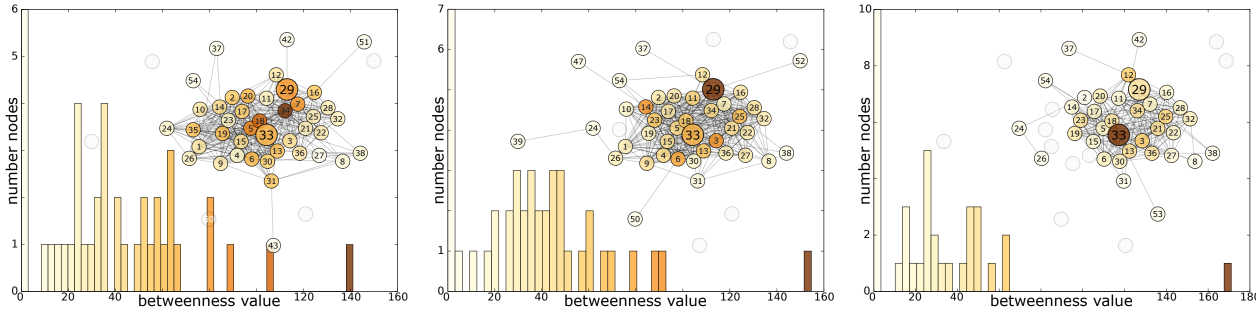
SimBet's shortcomings can therefore be summarized as follows:

1. Representing a social relation in a binary form is not realistic and will lead to inconsistencies when comparing such relations.

2. The macro structure of a social (dynamic) network is changing over time, i.e. nodes enter, leave or switch their neighborhood and so the social role will become a different one.

3. The betweenness value calculation is meant to consider only shortest paths, but social communication is not limited only to the shortest possible paths.

4. The global betweenness utility of a single node becomes less important the closer a message comes to the destination. If the message reaches the destination's neighborhood, destination betweenness can be more precise by taking only relevant directions into account.



(a) Using only the first third. (b) Using only the second third. (c) Using only the third third.

Figure 1: The distribution of the ego betweenness using the original SimBet algorithm processing slices of the *imote* trace. The structure changes over time. Nodes with a high betweenness in subfigure (b), e.g. node 29, have a very low betweenness at a later point in time, while others display a higher betweenness only later, e.g. node 33 in subfigure (c). The original SimBet algorithm does not reflect these changes when run on the complete data set.

3. SIMBETAGE

In this section, we present SimBetAge. While building upon SimBet, we deal with the dynamics of social networks and adjust the used social network metrics significantly. Doing so, we increase performance drastically.

3.1 Modelling the dynamics in relations

A binary model of a social relation, i.e. one node knows another node or not, does not cope with the dynamics such a network can have. Using a weighted time dependent graph $G(t)$ as a model of a social network will bring a more realistic view, where $G(t) = (V, E, \omega(e,t))$ is a fully connected graph. This weight of an edge is what we will call the *freshness* of an edge, where $\omega(e,t) = 0, e = (u,v) \in E, t \in T$ means, that the nodes u and v have never be connected from time t_0 up to time t and $\omega(e,t) = 1$ represents a permanent connection between u and v.

The *freshness* of a single contact is influenced by two events: The first one is a time step event that will decrease the freshness value, i.e. the contact becomes older, which is modeled through the use of an exponential decay function. The second one is an encounter event that will increase the freshness value, i.e. the contact becomes *fresher* which is modeled by a logistic growth function. The equations of the two events are shown in 4 and 5.

$$\omega_{\text{new}} = \omega_{\text{old}} \cdot \gamma^{\Delta t} \tag{4}$$

$$\omega_{\text{new}} = \omega_{\text{old}} + (1 - \omega_{\text{old}}) \cdot \alpha \tag{5}$$

Figure 2 shows the exemplary development of a freshness value over time.

After we have defined the freshness of a single edge, we also need to define the freshness of a path P with length $l(P) > 1$. Due to the representation by logistic growth and exponential decay for the age value of a single edge, the freshness $\omega(e,t)$ can also be interpreted as an indicator, how probable it is that the two nodes u and v are connected at time t. Hence, the freshness $\omega(P,t)$ of a path P is defined by the product of all freshness values in P.

$$\omega\left(P,t\right) := \prod_{e \in P} \omega\left(e,t\right) \tag{6}$$

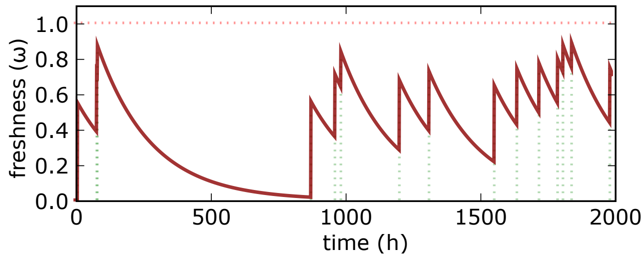


Figure 2: The *freshness* ω of a single edge with some encounter events over time. The higher the value, the greater the probability of encounters in the near future.

With these two definitions for aged edges and aged paths we now define our calculation for similarity and betweenness in aged graphs.

3.2 Similarity for Aged Graphs

In a binary graph the similarity $S_{\text{binary}}(u,v) = |N(u) \cap N(v)|$ of two nodes u and v is defined by the number of common neighbors between u and v. In a graph with aged relations, we define the similarity of two nodes as being proportional to the freshness of the concurrent neighborhood. If node u and node v share a neighbor w with the freshness values $\omega(u,w)$ and $\omega(v,w)$, then the similarity $\sigma_w(u,v)$ of node u and v with respect to neighbor w is asymptotically equal to the product of the freshness values of their relations to w.

The calculation of the similarity is based on the encounter information exchanged during runtime. Therefore, we need to consider that the encounter information at node u about node v might be outdated proportionally to the time node u has not heard about the encounters of node v with w either from node v or from node w. This means, the similarity of node u and node v concerning w needs to be rated by the freshness of the encounter information of node v about the common neighbor w.

$$\sigma_w(u,v) := \begin{cases} \omega_{u,w} \cdot \omega_{v,w} \cdot \omega_{u,v} & \text{if } v \in N_1(u) \\ \omega_{u,w}^2 \cdot \omega_{v,w} & \text{else} \end{cases} \tag{7}$$

Following this definition, we then define the aged similarity $\sigma(u,v)$ as the sum of the similarities of all common neighbors i between node u and v.

$$\sigma(u,v) := \sum_{w \in N_1(u) \cap N_1(v)} \sigma_w(u,v) \tag{8}$$

3.3 Betweenness for Aged Graphs

Analogous to the calculation of the betweenness in binary graphs, in weighted graphs only geodesics between two nodes are taken into account. In an aged graph, those are the paths with the highest freshness value $\omega(P,t)$. The method for calculating stays the same as the binary betweenness as in equation 2. However, the geodesics may change suddenly and significantly depending on the individual freshness values, while for routing decisions, smoother functions are preferable since they better model the connectivity in social networks.

3.4 Flow Betweenness

When looking at communication characteristics in social networks, the question arises why only shortest paths linking two individuals should be the possible way of communication flow. In a stochastic model of a network, you might be better advised to use all possible paths between two nodes for a more representative view.

According to this model, [7] defines the flow betweenness. Instead of the number of geodesics, the calculation takes all possible paths in a network into account, which was inspired by [14]. As a result of their observations they came up with a centrality measure more consistent and susceptible to dynamic network processes than other centrality measures.

Flow betweenness centrality is based on Ford and Fulkerson's [6] model of network flows. We therefore define the ego-centric flow betweenness value ϑ_u as the sum over the age of all paths between pairs of neighbors (v,w) of u passing through u divided by the age of all possible paths between them, weighted with the age of the edge between u and v and the one between u and w:

$$\vartheta_u := \sum_{v,w \in N_1(u)} \frac{(\omega_{u,v} \cdot \omega_{u,w})^2}{\omega_{u,v} + \sum\limits_{\substack{x \in \{u\} \cup N_1(u) \\ x \neq v \neq w}} \omega_{x,v} \cdot \omega_{x,w}} \tag{9}$$

Figure 3 demonstrates the influences of aging on the flow betweenness distribution of a network evolving over time. Comparing it with figure 1, we see that the distribution follows the dynamics and structural changes in the network now very closely, thus making the representation more consistent with the real world.

3.5 Directed Betweenness

In an ego-centric view we distinguish between three decision states when two nodes u and v compare their utilities for a destination d:

1. Both nodes are far away from the destination $d \notin N_1(u) \cup N_1(v)$
2. At least one node has the destination in its two hop proximity $d \in (N_2(u) \cup N_2(v)) \setminus (N_1(u) \cup N_1(v))$
3. At least one node is close to the destination $d \in N_1(u) \cup N_1(v)$

In the first case the decision can easily be reached using the betweenness value, while in the third case the higher similarity will be the best choice. However, the second case might be critical as the similarity might not be meaningful enough and a high betweenness could possibly lead in the wrong direction when reaching the proximity of the destination. Therefore, we introduce the directed betweenness measure as a slight modification to the general ego flow betweenness: Instead of all possible paths in the neighborhood of u, only those containing d are considered.

$$\delta_u(d) := \sum_{\substack{v,w \in N_1(u) \\ d \in N_1(v) \vee v = d}} \frac{(\omega_{u,v} \cdot \omega_{u,w})^2}{\omega_{u,v} + \sum\limits_{\substack{x \in \{u\} \cup N_1(u) \\ x \neq v \neq w}} \omega_{x,v} \cdot \omega_{x,w}} \tag{10}$$

3.6 Defining an Utility

Keeping in mind the considerations of the previous section, we define an utility for the node v with a message bundle for d currently in contact with node n. Depending on the neighbor relations we weigh the similarity $\sigma_v(d)$ strongest, the destination betweenness $\delta_v(d)$ second and the general betweenness ϑ_v weakest[1]:

$$A = \begin{cases} 0.9 & \text{if} \begin{cases} v \to d \text{ or} \\ v \to n \to d \end{cases} \\ 0.4 & \text{if} \begin{cases} v \to \not n \to d \text{ or} \\ n \to \not v \to d \end{cases} \\ 0 & \text{else} \end{cases} \tag{11}$$

$$B = \begin{cases} 1 - A & \text{if} \begin{cases} d \notin N_2(v) \text{ and} \\ d \notin N_2(n) \end{cases} \\ 0 & \text{else} \end{cases} \tag{12}$$

$$C = 1 - A - B \tag{13}$$

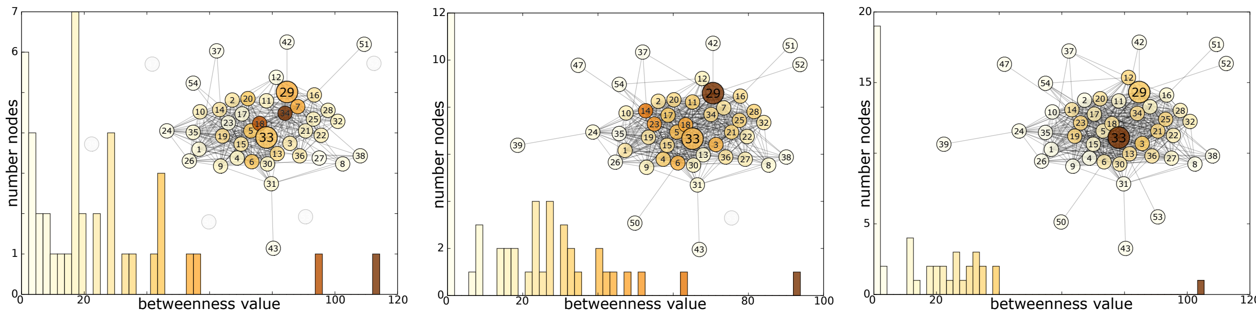
$$U_v(d) = A\sigma_v(d) + B\delta_v(d) + C\vartheta_v \tag{14}$$

3.7 Update Messages

In our reference implementation of the considered routing algorithms, routing decisions are made at the point in time a new encounter occurs. This may result in situations where a node will not forward a newly received bundle to a still connected and possibly better suited neighbor because of the absence of an encounter notification. In general, this problem should be solved by an underlying routing layer that transparently communicates in the temporarily connected subnetwork using a MANET algorithm. On the other hand, such layer is only useful if there is a meaningful number of connected subnetworks at any time.

We overcome these drawbacks by introducing an ***Up****date message* scheme involving a lookup to the neighbor table, and, if the node is connected to the destination, resulting in a direct delivery to the receiver. This way of updating does not influence the routing decisions or performance itself, but can be interpreted as a evidence on how well a routing performs on forwarding a bundle to the one hop proximity of the destination, and shows in which situations a subnetwork forwarding strategy might be useful.

[1]These constants were determined empirically. A general method to derive these constants is still under research.



(a) After the first third. (b) After two thirds. (c) After the whole trace.

Figure 3: Snapshots of the distribution of the ego betweenness using our algorithm for betweenness for aged graphs described in section 3.3 in the *imote* trace. It closely follows the change in structure in figure 1, e.g. in (b) node 29 has a high betweenness value, while in (c) node 33 has the highest betweenness value.

4. EVALUATION

As DTN routing algorithms are still a relatively young research area, best practices in evaluation are still emerging. In the following, we give a short overview over the metrics and trace data sources we used for evaluating SimBetAge.

4.1 Metrics

As DTN algorithms are used for communication, the arrival of data is a focal point. The *delivery rate* signifies how many messages out of a theoretical maximum arrive. The received messages may have taken widely different paths. The *path length* and the *delivery time* can be used to characterize these paths. Ideally, both should be minimized, but a path with more hops might result in faster delivery.

Additionally, depending on the algorithms used, different amounts of meta-data might be needed. This *control message overhead* expresses at which additional communication cost a certain performance can be achieved. This overhead might make the transfer of data messages prohibitively long or energy consuming, resulting in the need of careful evaluation in a given scenario.

4.2 Traces

Connectivity over time in delay tolerant networks depend heavily on the use case. Although simulation studies typically make use of random waypoint models (see [11] for an overview over methods in MANET simulation studies), we decided to use traces produced by actual experiments on social structures to properly capture the social behavior of participants. In the following, we elaborate on the traces used and on some of their characteristics.

In the context of the **MIT Reality mining** project [3], a group of 100 subjects at MIT over the course of the 2004/5 academic year. It is one of the most comprehensive records of social interactions available at present and was also used by the authors of the original SimBet paper [1].

In our experiments, we decided to use only the second half of these traces, as the first half includes only about 4% of the actual contact information. In the following, we call this trace *MIT50*.

Another project collecting traces of human interaction is the **Haggle** project [13]. A variety of users in office and conference environments collected Bluetooth sightings. We actually filtered these traces to only include sightings of subjects actively interacting with others, while filtering out passive devices like bluetooth headsets, laptops, and so on. In this paper, we call this trace *imote*.

Third, the traces generated in the **Dartmouth Outdoor experiment** [8] closely resemble a random waypoint simulation. A group of people were walking around an athletic field and collecting GPS information. In the beginning and in the end, all the subjects were meeting up at a central location, so that we actually used only the second and third quarter of this trace. In the following, we call this trace *darmouth*.

4.3 Results

In this section we compare seven different routing strategies. The first strategy is **Direct Delivery**. Only messages where the sender itself meets the recipient are delivered. This strategy is used as the base line as the minimum performance a strategy should achieve.

Epidemic Forwarding represents a second strategy, where every message is forwarded to anyone a node meets. This results in a significant message overhead. However, it serves as a theoretical upper bound, in terms of delivery rate.

Next, we consider the original **PRoPHET** algorithm, as described in [12] and the original **SimBet** algorithm as described in [1].

Finally we evaluate different variants or our proposed algorithm. The first variant **SimBetAge** replaces the binary representation of similarity and betweenness, but does not include the directed betweenness, introduced in section 3.5. The next variant, **Dest2SimBetAge**, follows the design exactly as presented in the SimBetAge section, including directed betweenness and strictly local knowledge. This is the variant we propose as the best trade-off. As a final variant, **DestSimBetAge**, we extend Dest2SimBetAge to consider all known transitive links for the directed betweenness calculation, similar to PRoPHET. This variant performs only slightly better then Dest2SimBetAge, but needs more state information.

Figure 4 compares the packet delivery rate achieved by the different routing strategies. While our performance is better in all traces considered, the most drastic improvement can be seen in the MIT50 trace, which is incidentally the longest running trace in which the most change happens. In this trace, the delivery rate is 95% better than SimBet and about 6% better than PRoPHET.

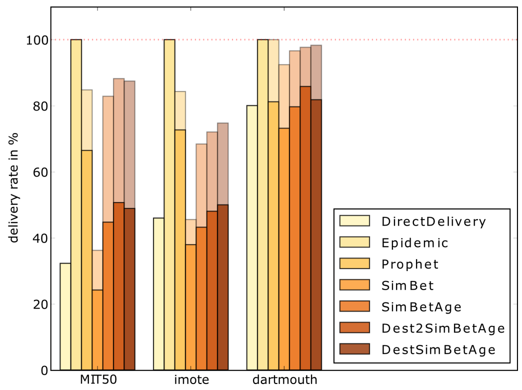


Figure 4: Bundle delivery rate for different routings.

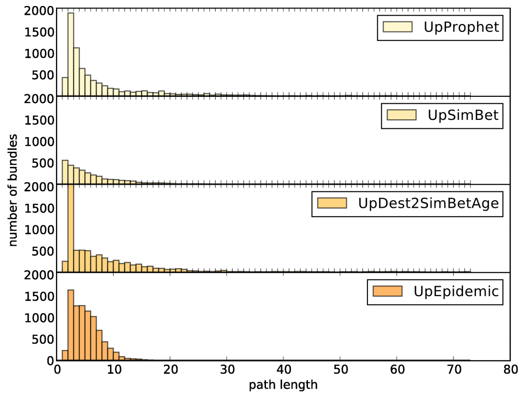


Figure 5: Bundle delivery path length for the MIT50 trace.

As can be seen in 5, the hop count distribution of our approach is significantly better compared to PRoPHET, while the delivery rate as shown in figure 6 stays comparable.

With respect to control message overhead and total message overhead, which are not shown here, PRoPHET and our approach are comparable, while we reduce the needed state information, and therefore bytes transferred, significantly, using only an ego-centric representation.

In summary, respecting the age of node relations leads to a better performance than a binary representation. The directed betweenness is a very useful addition to SimBetAge and performs better on all traces. However, the additional distance information introduced in DestSimBetAge does not affect the performance in a significant positive way and supports our decision on keeping the ego-centric view instead of producing more overhead on discovering nodes not covered by the two hop neighborhood.

5. CONCLUSIONS AND OUTLOOK

We presented our extensions to SimBet which also take changes in the social structure over time into account. For that purpose, we defined similarity as well as ego flow betweenness in weighted graphs. Furthermore, we evaluated these extensions for the purpose of forwarding messages in delay and disruption tolerant networks, specifically those based on social networks. We showed that the performance can be drastically improved when temporal changes are considered.

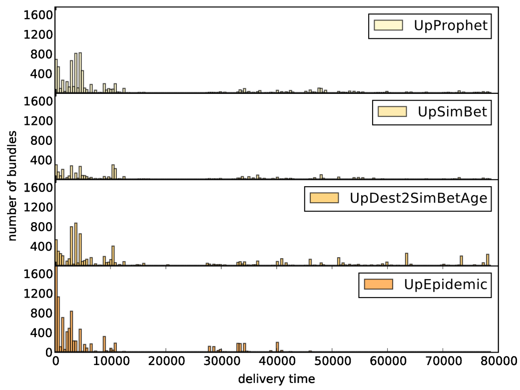


Figure 6: Bundle delivery time for the MIT50 trace.

In the near future, we plan to implement this algorithm in the context of the RatPack project. An interesting point we still left open in this paper is the optimal choice of decay and growth parameters as well as how to best mitigate discretization artifacts in the context of implementations in embedded systems.

6. REFERENCES

[1] E. M. Daly and M. Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *MobiHoc '07*, Montreal, Quebec, Canada, 2007. ACM Press.

[2] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli. Age matters: efficient route discovery in mobile ad hoc networks using encounter ages. In *MobiHoc '03*, New York, NY, USA, 2003. ACM Press.

[3] N. Eagle and A. S. Pentland. CRAWDAD data set mit/reality (v. 2005-07-01). Downloaded from http://crawdad.cs.dartmouth.edu/mit/reality.

[4] M. Everett and S. P. Borgatti. Ego network betweenness. *Social Networks*, 27(1):31–38, 2005.

[5] S. Farrell and V. Cahill. *Delay- and Disruption-Tolerant Networking*. Artech House Publishers, September 2006.

[6] L. Ford and D. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8(3):399–404, 1956.

[7] L. C. Freeman, S. P. Borgatti, and D. R. White. Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, 13(2), June 1991.

[8] R. S. Gray, D. Kotz, C. Newport, N. Dubrovsky, A. Fiske, J. Liu, C. Masone, S. Mcgrath, and Y. Yuan. CRAWDAD data set dartmouth/outdoor (v. 2006-11-06). Downloaded from http://crawdad.cs.dartmouth.edu/dartmouth/outdoor.

[9] P. Hui, A. Chaintreau, R. Gass, J. Scott, J. Crowcroft, and C. Diot. Pocket Switched Networking: Challenges, feasibility and implementation issues. *LNCS*, 3854:1, 2006.

[10] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: social-based forwarding in delay tolerant networks. In *MobiHoc '08*, New York, NY, USA, 2008. ACM Press.

[11] S. Kurkowski, T. Camp, and M. Colagrosso. Manet simulation studies: the incredibles. *SIGMOBILE Mob. Comput. Commun. Rev.*, 9(4):50–61, October 2005.

[12] A. Lindgren, A. Doria, and O. Schelén. Probabilistic routing in intermittently connected networks. In *Service Assurance with Partial and Intermittent Resources*, volume 3126/2004 of *LNCS*, pages 239–254. Springer, 2004.

[13] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau. CRAWDAD data set cambridge/haggle (v. 2006-09-15). Downloaded from http://crawdad.cs.dartmouth.edu/cambridge/haggle.

[14] K. Stephenson and M. Zelen. Rethinking centrality: Methods and examples. *Social Networks*, 11, 1989.