On the Strength of Weak Ties in Mobile Social Networks

Stratis Ioannidis
Thomson, Paris
stratis.ioannidis@thomson.net

Augustin Chaintreau
Thomson, Paris
augustin.chaintreau@thomson.net

ABSTRACT

Weak ties, introduced in the seminal paper by Granovetter [6], refer to people's relationships with acquaintances outside their social cirle (*i.e.*, acquaintances with which they interact infrequently). We show that weak ties play an important role on the dissemination of content updates over a mobile social network. From a practical perspective, weak ties can be used to design strategies for limiting the bandwidth usage on such a network, without severely impacting the speed with which the content is disseminated. Moreover, the importance of weak ties can be traced on the effect they have on the edge expansion of the social network. By identifying this relationship, we extend and validate earlier work [7], that characterizes the speed of content dissemination in terms of the network's edge expansion.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Network topology

General Terms

Measurement, Performance

Keywords

Online social networks, mobile networks, delay-tolerant networks, pocket-switched networks, edge expansion

1. INTRODUCTION

The proliferation of mobile phones and other portable devices with increased storage and wireless communication capabilities has created a new networking environment for peer-to-peer applications. Mobile users can take advantage of encounters with other users to share content (*e.g.*, using Bluetooth) in a peer-to-peer fashion. Interestingly, such opportunistic exchanges of content depend on the users' social behavior in the following two ways. First, people meet more frequently, and more regularly, with people within their social circle (*e.g.*, their family, friends and co-workers). Second, their willingness to share content may depend on their social ties, as one may be reluctant to trust content originating from a stranger. As a result, it has been observed [9, 11, 4, 7] that the performance of a peer-to-peer application that relies on opportunistic exchanges depends on the *social network* formed by the mobile users, determined by the relationships and interactions between them [8].

This dependence is clearly illustrated by the following content-distribution application, first proposed and analyzed in [7]. In this application, mobile users subscribe to a dynamic-content distribution service, offered by their wireless service provider. For example, users might receive updates from a news-feed or a blog, a traffic congestion monitoring service, *etc.* In addition, users that subscribe to this service share their downloaded content with each other in an opportunistic, peer-to-peer fashion: whenever two subscribers meet, the one whose content is most recent pushes it to the one whose content is older, thereby extending the coverage of the service and increasing the network's capacity.

The information provided to subscribers should be as "fresh" as possible; for this reason, an important measure of the quality of this service is the age of the content available to a subscriber. In general, the more frequently users meet and share their updates with each other, the fresher the content in the system will be. In this sense it is preferable to exploit as many opportunities for sharing content as possible in order to improve the performance of the content-distribution network. On the other hand, subscribers may not wish to share content every time they meet another subscriber, *e.g.*, to limit their bandwidth usage and reduce their battery consumption. Hence, there is an inherent trade-off between the bandwidth used by the application and the

SNS'09, March 31, 2009, Nuremberg, Germany.

performance of the system, as captured by the age of the content available to subscribers.

In general, it is not a priori clear when a user should forego the opportunity to share content. For example, should users limit their bandwidth usage by foregoing contacts with people they meet very often, or with people they meet rarely? To answer this question, some intuition can be gained from the following classic notion from sociology, introduced in the seminal work by Granovetter [6]. Granovetter argues that we can separate our relationships into *strong ties*, *i.e.*, relationships with people in our immediate social circle, and *weak ties*, *i.e.*, relationships with acquaintances that we have little interactions with. Strong ties lead to the creation of closely knit, clustered communities (*i.e.*, cliques, in graph-theoretic terms) in which everyone is connected to each other. On the other hand, weak ties play the important role of bridging such communities together. In this sense, weak ties can be very beneficial to individuals, as they provide access to communities they do not belong to.

The first contribution of this paper is to show that weak ties play an important role in how content is disseminated in our system. In particular, the strategy employed by users to limit their bandwidth usage should *preserve weak ties*: users should forego contacts with people they meet very often and utilize contacts with people they meet rarely. By doing so, bandwidth usage can be reduced significantly without severely impacting the age of the content available at users. We obtain the above result by studying the trade-off between bandwidth usage and content age achieved by simulating several different strategies over three different datasets of human mobility traces.

The second contribution of this paper is to show that weak ties are important for content dissemination precisely because of their effect on a graph-theoretic property of the social network, namely, its edge expansion. By doing so, we extend and validate the analysis of [7] that characterized the content age in terms of the social network's edge expansion. In particular, in this paper we show that, instead of focusing on the content age, one can compare different strategies for limiting the bandwidth usage by investigating their effect in the edge expansion. This result is important for the following reasons. First, it suggests a less computationally intensive alternative method for evaluating the performance of different strategies for limiting bandwidth usage. Second, it gives another explanation why weak ties are so important: by connecting otherwise disjoint clusters, weak ties have the effect of improving the expansion of the social network and, indirectly, increasing the speed with which content updates are disseminated. Finally, it validates the relationship between edge expansion and content age, as predicted by [7].

The remainder of this paper is structured as follows. In Section 2, we review the related work in the area. In Section 3, we give a more detailed description of the system we analyze, and in Section 4 we present our analysis based on human mobility traces. Finally, we conclude by discussing possible future directions of this work in Section 5.

2. RELATED WORK

The importance of weak ties in the context of mobile social networks was first identified by Miklas *et al.* [9]. The authors observe experimentally that although the majority of contacts occur within clusters of friends, weak ties play an important role in keeping the social network connected. The authors further show that such connections can be exploited to improve the performance of applications like routing, propagating queries in DTNs and blocking the spread of viruses. In this sense, our work expands on [9] by identifying an additional application that can benefit considerably from the use of weak ties, as well as through explaining the importance of weak ties through their effect on the edge expansion of the social network.

Several additional algorithms taking advantage of the social network formed by mobile users have been proposed for publish/subscribe systems [11, 4], routing [1] and query propagation [10]. These algorithms exploit node centrality [11, 10], friendship relationships [11] and metrics of contact usefulness [1, 4]. Again, our work differs both in the application considered, as well as in identifying edge expansion (and, through it, weak ties) as the property of the social network that affects this application's performance.

3. SYSTEM DESCRIPTION

3.1 Content Update Process

We adopt the model presented in [7] for our analysis. The system we describe consists of n mobile users that are served by a single service provider (*e.g.*, a cellphone operator). We denote by $V = \{1, 2, \ldots, n\}$ the set of all users. The service provider pushes new content to user $i \in V$ according to a Poisson process with rate μ/n, where μ is service provider's downlink capacity.

The content stored at a user is time-stamped with the time at which it was originally downloaded from the service provider. We denote by $ts_i(t)$ the time-stamp of the content stored at user i at time t.

We say that a user $i \in V$ *contacts* or *meets* another user $j \in V$ when i and j are within each other's transmission range. Contacts are symmetric, *i.e.*, j contacts i whenever i contacts j, and they last for a time that is negligible compared to the *inter-contact* time, namely the time between two consecutive contacts.

Users share the content updates they have downloaded from the service provider with other users they meet. In particular, whenever two users i and j meet, they may or may not share their content updates. This will depend on the strategy employed by users to limit their bandwidth usage. We describe such strategies in more detail in Section 4. If the users decide to *not forego* the opportunity to share their updates, user i will copy user j's content if the content stored at j is more recent than the content stored at i, *i.e.*, $ts_i < ts_j$. As a result, after i and j have shared an update, the time-stamp at both becomes $\max(ts_i, ts_j)$.

The *age* of the content stored at user $i \in V$ at time t, denoted by $Y_i(t)$, is the following random variable:

$$Y_i(t) = t - ts_i(t), \qquad i \in V.$$

Our goal is to study the maximum expected age over all users when the system is in steady state. More formally, let

$$\mathbb{E}[Y_i] = \lim_{t\to\infty} \mathbb{E}[Y_i(t)], \qquad i \in V,$$

be the expected age of the content available at user i in steady state (*i.e.*, after the system has operated for a sufficiently long time). Note that by ergodicity, for all $i \in V$,

$$\mathbb{E}[Y_i] = \lim_{t\to\infty} \frac{1}{t} \cdot \int_0^t Y_i(s)ds, \quad a.s.,$$

i.e., $\mathbb{E}[Y_i]$ is also the time-average of the age of the content stored at user i. In our analysis, we characterize

$$\max_{i\in V} \mathbb{E}[Y_i]$$

i.e., the maximum expected age (in steady state), taken over all possible users. Intuitively, by considering the maximum expected age, we are capturing the "worst-case" behavior of the system.

3.2 Contact Process and the Contact Graph

As in the second part of [7], we make the simplifying assumption that the inter-contact time between any two users $i, j \in V$ is distributed exponentially. Hence, user i contacts another user j according to a Poisson process with rate $q_{ij} \geq 0$, which we call the *contact process* between i and j. In addition, we assume that the contact processes describing contacts between distinct pairs (i,j) are independent. Note that $q_{ij} = q_{ji}$, because contacts are symmetric.

Again, as in [7], we represent the above system with a complete, weighted, undirected graph $G(V,E)$ whose vertex set is the set V of mobile users and the weight of an edge (i,j) is $q_{ij} \geq 0$. We refer to $G(V,E)$ as the *contact graph* [11, 7] of the system.

We denote by $q_i = \sum_{j\neq i} q_{ij}$ the aggregate rate with which i contacts other users. Moreover, we denote by $q_{\max} = \max q_i$ and $q_{\min} = \min q_i$ the aggregate contact rates of the "most social"and the "least social" user in the system, respectively. Note that the number of content exchanges per time unit a user i engages in (*i.e.*, the bandwidth consumed by such exchanges on user i) is bounded by q_i.

3.3 Edge Expansion

A property of the contact graph that will play an important role in our analysis is its *edge expansion* [3]. Given a contact graph $G(V,E)$, its edge expansion is

$$h_G = \min_{A\subset V, |A|\leq \frac{n}{2}} \frac{\sum_{i\in A}\sum_{j\in V\setminus A} q_{ij}}{|A|}. \qquad (1)$$

Intuitively, for A a set of users, the quantity in the numerator in (1) is the "capacity" of the outgoing links of A: it is an upper bound on the rate with which content updates can exit the set. Moreover, recall that the service provider injects content updates in set A with rate $|A| \cdot \frac{\mu}{n}$, *i.e.*, proportional to $|A|$. The ratio in (1) therefore captures how quickly updates can exit a set, in proportion to the rate with which they are injected from the service provider. Hence, as h_G is a lower bound on this quantity, it captures the "worst-case" exit capacity from a set, in proportion to its down-link update traffic.

As such, it is not surprising that the edge expansion of the contact graph plays an important role in how quickly content is disseminated. Indeed, in [7], it is shown that the edge expansion can be related to the maximum content age over all users through the following bound [7, Theorem 2]:

$$\max_{i\in V} \mathbb{E}[Y_i] \leq \frac{2}{\mu}\left(2e^{-1/2} + \log(n)\right) + h_G^{-1}\log n\,. \qquad (2)$$

Computing the edge expansion of a given graph is an NP-hard problem. However, the edge expansion is closely related to the graph's spectral gap, which can be computed in polynomial time. In particular, let $\mathcal{L} = [l_{i,j}]_{i,j\in V}$, where

$$l_{ij} = \begin{cases} 1, & \text{if } i = j \\ \frac{q_{ij}}{\sqrt{q_i q_j}} & \text{o.w.} \end{cases}$$

Matrix $\mathcal{L}$ is known as the graph's Laplacian [3]. We then have that [3]

$$q_{\min}\lambda/2 \leq h_G < q_{\max}\sqrt{2\lambda}, \qquad (3)$$

where λ the second smallest eigenvalue of $\mathcal{L}$. We use the lower bound of (3) to obtain a pessimistic estimate of the edge expansion in our analysis of Section 4.

4. THE STRENGTH OF WEAK TIES

4.1 Content Age and Expansion on 3 Datasets

Our results are based on human mobility traces from three datasets collected by other researchers. Infocom06 [2] contains Bluetooth contacts between iMotes distributed to 78 participants of a three-day conference. We include in our analysis only the 65 participants for which we had sociological data (see Section 4.3). CoNext07 contains Bluetooth contacts observed between cellphones distributed to 27 participants in a similar three-day conference environment. The MIT dataset [5], consists of 95 participants carrying GSM cell-phones over a period of 9 months. As in [2, 7], we assume that two phones are in contact when they share the same GSM base station. We exclude 8 nodes from our analysis, as they are isolated.

For each of the three datasets, we calculate the rate of opportunistic contacts q_{ij} between a pair of nodes i, j as the inverse of the average inter-contact time seen during the experiment. We use the lower bound of (3) to evaluate h_G. Note that our evaluation can be used to bound the maximum content age through the inequality (2).

Let us first briefly consider the case where all contacts between nodes are used for exchanging content updates. For the values of μ indicated in Table 1, we computed through simulations the steady-state expected age (average and maximum over all nodes) obtained with and without content sharing between users. In Infocom06 the age of the content in a node drops from nearly two hours to a little more than 8 minutes, roughly ten times smaller, when users share their content. In this case, a node contacts other users every minute, on average. In the MIT dataset, the expected age drops from above one day to under 3 hours, while a contact occurs on average every 5 minutes. The bound of (2) gives a rough estimate for the Infocom06 and CoNext07 datasets, three and six times larger than the actual maximum expected age. The estimate of the MIT dataset is considerably less accurate, as the maximum expected age is overestimated by a factor of 90 (see also the discussion in the next section).

Data set	Infocom06	CoNext07	MIT
System size n	65	27	87
Duration (days)	3	4	246
μ (sec^{-1})	0.01	0.01	0.001
Node contact rate (sec^{-1})			
Maximum	0.0632	0.0104	0.0170
Average	0.016	0.0047	0.003
Minimum	0.0067	0.0028	0.00014
Expected age (sec)			
no sharing	6500	2700	87,000
sharing avg.	345±10	649±24	3,740±130
sharing max.	447±14	774±71	9,840±470
Bound of (2)	2,935	5,129	881,110

Table 1: Characteristics of the three datasets. Sharing content can significantly reduce both the average and the maximum age over all users.

In the rest of this section, we assume that nodes reduce their contact rates according to different strategies. We study how this affects both the maximum expected age obtained in steady state, as well as the edge expansion.

4.2 Bandwidth Usage Limitation Strategies based on Contact Rates

In this section, although the actual contact rate between two users i and j may be q_{ij}, a user may choose to utilize only a fraction $q'_{ij} \leq q_{ij}$ of its available contacts. We assume that users reduce their bandwidth usage according to the following strategies, in which the decision to forego certain opportunities for sharing content is based only on the contact rates q_{ij}.

Keep-Highest(k) [KH(p)]. Every user i restricts its contacts to only k users. These k users are the ones with the highest contact rates with this node.

Proportional(p) [P(p)]. Every contact rate is scaled by $0 < p < 1$, *i.e.*, $q'_{ij} = pq_{ij}$. This can be implemented by ignoring contact events with probability $1 - p$ and using them for content sharing with probability p.

Proportional-Keep-Highest(p) [PKH(p)]. A user reduces its aggregate contact rate to pq_i, $0 < p < 1$, by removing a rate $(1 - p)q_i$ in total from its neighbors in the contact graph with which it has the lowest contact rate.

Proportional-Keep-Lowest(p) [PKL(p)]. A user reduces its aggregate contact rate to pq_i, $0 < p < 1$, by removing a rate $(1-p)q_i$ in total from its neighbors in the contact graph with which it has the highest contact rate.

Rate-Limit (q) [RL(q)]. For a given q, each user i with $q_i > q$ equalizes its highest contact rates (in other words, trims them), so that its aggregate rate is $q'_i = q$.

The last four strategies for user i are illustrated in Figure 1, where nodes are sorted according to their contact rates with i. The strategy Proportional(p) is neutral, as it does not change the relative strength of any edge in the network. All other strategies discriminate between strong ties (edges with high contact rates q_{ij}) and weak ties (edges with low contact rates q_{ij}). Moreover, with the exception of Proportional(p), all strategies may yield asymmetrical rates: For example, if j is among i's k highest rate contacts while i is not among j's k highest rate contacts then Keep Highest(k) gives $q'_{ij} = q_{ij}$ and $q'_{ji} = 0$. To address this, we consider two variants of each of the above strategies. In the first, called Keep-If-*Any*-Keep (denoted by the suffix "Any" in Figures 2 to 5), an exchange occurs as long as *any* of the two users wishes to share content. Equivalently, the reduced contact rate between these nodes is given by the

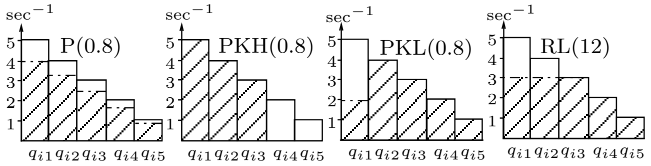

Figure 1: The Proportional(p), Proportional-Keep-Highest(p), Proportional-Keep-Lowest(p), and Rate-Limit (q) strategies. The unshaded regions indicate the contact rates q_{ij} (in contacts per second). The shaded regions indicate the resulting contact rates q'_{ij} after a bandwidth-usage reduction strategy is applied.

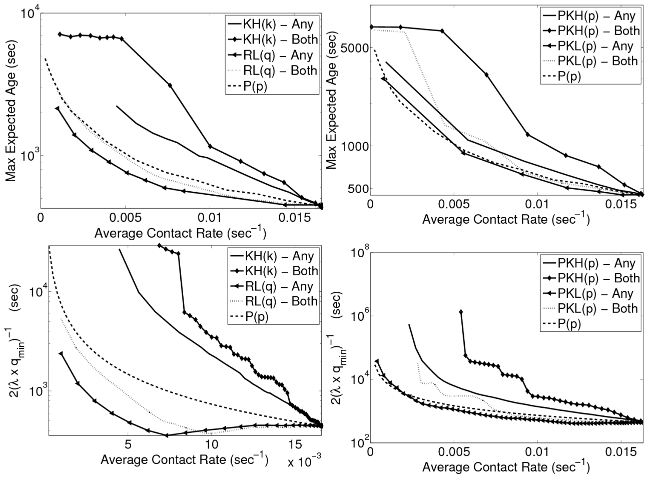

Figure 2: Maximum content age and edge expansion for different bandwidth-usage limitation strategies in the Infocom06 dataset. The edge expansion, shown in the bottom row, is approximated by $\frac{\lambda q_{\min}}{2}$ (see (3)). The most favorable trade-offs between maximum content age and bandwidth usage are achieved over strategies that preserve weak ties, such as PKL(p) and RL(q). Moreover, the qualitative comparison between the strategies in terms of the inverse of the edge expansion agrees with the one derived through the maximum age.

maximum of q'_{ij} and q'_{ji}. In the second variant, called Keep-If-*Both*-Keep (denoted by "Both" in Figures 2 to 5), content is shared up to a rate allowed by *both* nodes; the reduced contact rate is then $\min(q'_{ij}, q'_{ji})$.

We study the trade-off established by these strategies between the bandwidth used in the network and the content age. The top row of figures 2 and 3 present the maximum age per user versus the average contact rate $\sum_{i \in V} q'_i/n$ under each of the above strategies, for the Infocom06 and MIT datasets, respectively.

First, we see that reducing bandwidth is possible without necessarily leading to high content ages. For instance, the Proportional(p) strategy can be used to reduce bandwidth by half, while keeping an age under 700sec (resp. 11000sec) in Infocom06 (resp. MIT). All our results agree that the best performance trade-off is achieved by strategies that prioritize weak over strong ties. In particular, strategies may be sorted in decreasing order of efficiency as Rate-Limit, Proportional-Keep-Lowest, Proportional, Proportional-Keep-Highest and Keep-Highest. In other words, removing bandwidth from strong ties can reduce the bandwidth consumption considerably while incurring a lower age than the one experienced at the same average bandwidth when removing weak ties. With the Rate Limit strategy, that achieves the best trade-off in both datasets, one can remove 75% of the contacts without even doubling the maximum content age. We note, in addition, that the variant Keep-if-Any-Keep of every strategy is consistently better than the Keep-if-Both-Keep variant; this can be attributed to the fact that the latter penalizes users with smaller contact rates.

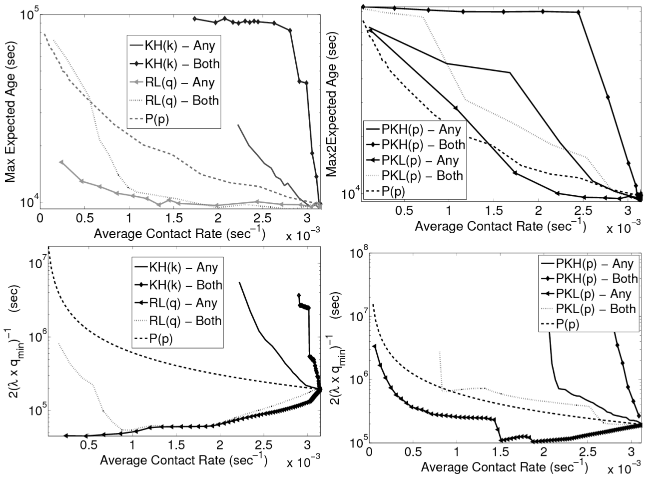

Figure 3: Maximum content age and edge expansion for different bandwidth-usage limitation strategies in the MIT dataset. The observations made on the Infocom06 dataset (Figure 2) also apply here.

In the bottom row of Figures 2 and 3, we plot an estimate of h_G^{-1}, for the contact graph with weights q'_{ij}. This estimate is obtained by the lower bound of Inequality (3), and is again plotted as a function of the average contact rate. We observe that the qualitative comparison between the strategies in terms of h_G^{-1} agrees with the one derived through the maximum content age for both datasets. This is quite surprising, given that Eq. (2) merely establishes an upper bound on the maximum age, and that (3) overestimates h_G^{-1}. This suggests that, despite the approximations involved, the age of the system is inherently linked to the edge expansion of the contact graph. For Rate Limit(q), we see that the bound we obtain becomes smaller as more contacts are removed, especially for the MIT dataset. In contrast, the age can only increase as contacts are removed. This simply indicates that the approximation made for h_G in Eq.(3) becomes more accurate, which also explains the gap that we observed for the MIT dataset in Table 1.

4.3 Bandwidth Usage Limitation Strategies based on Social Preferences

Users may decide to forego opportunities to share content based on trust issues. In particular, cooperation between users that belong to the same social circle is more likely; for this reason, it is interesting to understand the importance of weak ties in this context as well. To do so, in this section, we investigate the effect of limitations on bandwidth-usage that rely on social preferences.

For the Infocom06 and CoNext07 data-sets we could access sociological data, which we used to define social relationships between the participants. In Infocom06, participants filled out a questionnaire and indicated their technical interests from the 35 topics announced in the call for papers of the conference (*e.g.*, network architecture, power control *etc.*). In our analysis, we assume that people sharing a larger number of topics of interest are closer in a social sense. In CoNext07, participants explicitly declared their friends, which allowed us to create a graph representing the friendship relationships between them. We assume that people that have a smaller distance in the friendship graph are closer in a social sense. Note that two friends are at distance 1, and that two users with a common friend are within distance at most 2.

We consider the following strategies:

Keep-Friends(k) [KF(k)]. User i restricts its contacts to only its k closest users, based either on the number of shared interests or their distance in the friendship graph.

Keep-Strangers(k) [KS(k)]. User i restricts its contacts to only k other users, chosen as the furthest ones in a social sense. This is the exact opposite of the previous strategy.

Rate-Limit-Keep-Friends (resp. Strangers)(q) [RLKF(q)-RLKS(q)].
Every user with contact rate above q reduces its contact rate to q. The dropped rate is removed from the nodes furthest (resp. closest) to the user.

As in the previous section, these strategies may create asymmetrical rates q'_{ij} on the edges of the contact graph. Again, two variants are considered: Keep-if-Any-Keep and Keep-if-Both-Keep.

As before, we study the efficiency of these strategies in terms of the age versus bandwidth consumption trade-off. We include Proportional(p) in our figures, as a neutral, reference strategy. We only show the Keep-if-Any-Keep variant, which was consistently better than the Keep-if-Both-Keep variant.

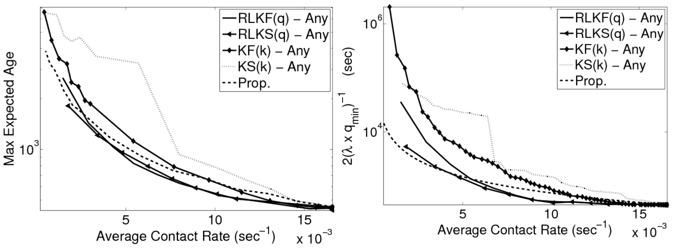

Figure 4: Infocom06: Maximum content age and edge expansion expansion for interest-based strategies.

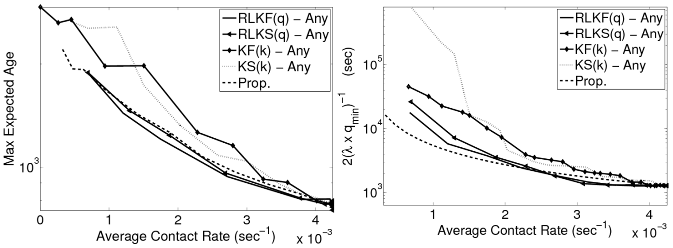

Figure 5: CoNext07: Maximum content age and edge expansion for friendship-based strategies.

As seen in Figures 4 and 5, contrary to contact-rate based strategies, we do not observe a clear differentiation between strategies that prioritize weak over strong ties. This can be attributed to the limited scope of the sociological data available to us. Nonetheless, we can conclude that the Rate Limit strategy outperforms the other two strategies in both datasets. Moreover, removing contacts according to a socially-based ordering, as opposed to simply the contact rate rank, does not significantly penalize this strategy. It remains close to Proportional(p) and occasionally provides a better trade-off. Most importantly, all qualitative comparisons of trade-offs between different strategies are correctly predicted by the study of edge expansions.

5. CONCLUSIONS

Our results have two important implications for the design of content dissemination systems over mobile social networks. First, weak ties play an important role in the dissemination; for this reason, when reducing bandwidth usage, preserving weak ties can achieve favorable bandwidth vs. performance trade-offs. Second, the importance of weak ties can be attributed to their effect on the network's edge expansion. In other words, we show that to capture the complex relationship between the social network formed by users and the performance of our distributed application it is sufficient to evaluate the edge expansion of the contact graph modelling user communication. The latter approach is both more insightful and less computationally intensive than simulating the system.

The importance of weak ties and, more generally, the impact of the contact topology are both closely related to key challenges for future distributed applications deployed on top of mobile social networks. Our results point to a few promising directions of research to address the following issues:

- **Finding better bounds:** The current definition of the edge expansion covers the worst case seen among all users as well as during each step of the dissemination process. It appears sufficient for many cases, especially to conclude on a qualitative comparison. On the other hand, occasionally it can be overly pessimistic. It seems both possible and promising to improve this bound towards an accurate evaluation of content age in heterogeneous networks.

- **Designing better strategies:** The bandwidth usage reduction strategies considered in this paper are fairly simple; nonetheless, they achieve good trade-offs between content age and bandwidth usage. Investigating more sophisticated mechanisms is an interesting open question; the notions of weak ties and edge expansion can again be helpful in understanding their behavior. One could even think of designing from scratch a strategy that precisely aims at maximizing the edge expansion. Moreover, we observe an important difference between strategies depending on the way asymmetry is handled, which indicates that reciprocation between users may play an important role.

- **Assessing the impact of inter-contact times statistics:** Our work on edge expansion, motivated by a known bound on the maximum expected age, focuses on exponential inter-contact times. However, the work in [7] suggests that dissemination is amenable to analysis even under more general processes.

- **Large-scale validation:** Most of our results are obtained for a population of nodes within a single large event or a relatively large community (multi-track conference, people on a university campus). We are not aware of data sets publicly available that could be used to extend this analysis to a large population. Such an extension may become possible as traces from mobile phone networks are released. In a larger population, the importance of weak ties may be even stronger, considering that they may connect two communities that are otherwise isolated. On the other hand, weak ties may also be too infrequent to be useful. This point, again, can be investigated through a similar study. Moreover, the preliminary results we obtained on strategies based on social preferences indicate that there may not be a simple correspondence between Friends (resp. Strangers) and Strong ties (resp. Weak ties). This point also requires further study.

6. ACKNOWLEDGMENTS

We would like to gratefully acknowledge Jason LeBrun, Earl Oliver and Anna Kaisa Pietilainen, for sharing their data with us, and Eran Gabber, for carefully reviewing this paper and helping improve its presentation.

7. REFERENCES

[1] A. Balasubramanian, B. Levine, and A. Venkataramani. DTN routing as a resource allocation problem. In *Proc. of ACM SIGCOMM*, 2007.

[2] A. Chaintreau, A. Mtibaa, L. Massoulié, and C. Diot. The diameter of opportunistic mobile networks. In *Proc. of ACM CoNEXT*, 2007.

[3] F. R. K. Chung. *Spectral Graph Theory.* AMS, 1997.

[4] P. Costa, C. Mascolo, M. Musolesi, and G.-P. Picco. Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks. *In IEEE JSAC Special issue on Delay-Tolerant Networks*, 2008.

[5] N. Eagle and A. Pentland. Reality mining: Sensing complex social systems. *Journal of Personal and Ubiquitous Computing*, 10(4):255–268, 2005.

[6] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[7] S. Ioannidis, A. Chaintreau, and L. Massoulié. Optimal and scalable distribution of content updates over a mobile social network. In *Proc. of INFOCOM*, 2009.

[8] D. Kempe, J. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proc. of ACM KDD*, 2003.

[9] A. G. Miklas, K. K. Gollu, K. K. W. Chan, S. Saroiu, P. K. Gummadi, and E. de Lara. Exploiting social interactions in mobile systems. In *Proc. of Ubicomp*, 2007.

[10] V. Srinivasan, M. Motani, and W. T. Ooi. Analysis and implications of student contact patterns derived from campus schedules. In *Proc. of ACM MobiCom*, 2006.

[11] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft. A socio-aware overlay for publish/subscribe communication in delay tolerant networks. In *Proc. of ACM MSWiM*, 2007.