

# On the Distribution of Content Updates over a Mobile Social Network

Contestant: Stratis Ioannidis

joint work with Augustin Chaintreau and Laurent Massoulié

## Problem & Motivation

Our work [1,2] focuses on a content distribution system over a mobile network. In this system, users subscribe to a content distribution service offered by their wireless service provider, as shown in [Illustration 1](#). The content distributed is dynamic and, as a result, subscribers need to continuously receive fresh updates for this content. For example, cell-phone users may receive updates regarding a news-feed or a blog of their interest, the price of a stock, traffic information, etc.



Illustration 1: A dynamic-content distribution system. The subscribers receive updates regarding a news-feed, a blog of their interest, the price of a stock, traffic information, etc.

Users that subscribe to this service share their 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. There are several benefits from exploiting such opportunistic exchanges among subscribers. These exchanges can be used to extend the network's coverage and improve the service provided to any user whose access to the wireless network is intermittent. Most importantly, such exchanges can increase the network's capacity. In particular, by utilizing the bandwidth of connections between users, the service provider can support more subscribers at a lower cost.

A question arising in the above setting is how should the service provider allocate its downlink capacity to ensure that the content at users is as “fresh” as possible. For example, should it allocate its available bandwidth uniformly among subscribers? Alternatively, should it provide more frequent updates to the “most social” subscribers, i.e., the ones that meet other subscribers most often, in the hope that they would spread the content faster?

In general, the answer depends on the provider's downlink bandwidth as well as on how recent the content at users is required to be. Most importantly, it also depends on the users' social behavior, since the latter determines how they meet. For this reason, answering the above question requires us to understand how the social network formed by mobile users affects the performance of our application.

A second question of importance is how such a system scales as the number of users grows. If more subscribers are added to the system, will, e.g., the average age of content at users increase, thus degrading the service, and, if so, by how much? This too depends on the topology of the social network formed by the users. Ideally, one wishes to find a general condition under which the content age increases slowly as the network grows.

Our work addresses these two questions, by (a) finding the optimal way to allocate the service provider's downlink capacity and (b) determining the system's scalability under user population growth.

## Background & Related Work

Computing the optimal allocation of the service provider's downlink capacity has a strong relationship to problems of maximizing influence in traditional social networks. One such problem is marketing a product to a subset of potential consumers to insure its widespread adoption [3]. Another is determining which blogs one should read to quickly detect the outbreak of an important story [4]. The feasible solutions of both of these problems lie in a discrete domain (i.e., are sets of consumers or blogs, respectively). It has been shown, by Kempe et al [3] and Leskovic et al [4], respectively, that greedy algorithms find solutions within a constant approximation of the optimal for these two problems, in polynomial time. Our problem is more natural in a continuous domain (as we are dealing with rates). It is interesting to note that, if the service provider can only forward updates to a subset of all users, at equal rate, the resulting problem (restricted now to a discrete domain) can also be solved by a greedy algorithm within a constant approximation of the optimal.

Distributing content over a DTN has recently been proposed, independently from our work, by Altman et al. [5] and Gao et al. [6]. In both works, it is assumed that the process determining how content is injected in the system is fixed; instead, users try to make optimal decisions on whether to share their content or not (e.g., to reduce energy consumption). We further note that, in both works, it is assumed that the inter-contact times between users are independent and exponentially distributed.

Algorithms taking advantage of the social behavior of mobile users have been proposed in the context of DTNs for a variety of applications different than the content distribution we consider here. Examples include publish/subscribe systems [7] and routing [8]. Typically, these algorithms exploit concepts derived from sociology and social networks, such as node centrality [7] and strength of relationships [8].

## Uniqueness of Our Approach

To the best of our knowledge, our work is the first to address the issues of optimality and scalability of content distribution over a mobile social network. The optimization problem we solve consists a new type of “influence maximization” problem, as a mobile network differs considerably from previous models of traditional social networks [3]. Most importantly, working in a continuous domain allows us to actually find the optimal allocation of the downlink rate, rather than a solution that is within a constant approximation factor to the optimal, which was the case in [3] and [4].

Our approach is orthogonal to the approach of Altman et al [5] and Gao et al. [6], as it focuses on optimizing the process according to which updates are injected in the system. Moreover, although we too assume independent Poisson contacts for our result on scalability ([Theorem 2](#), below), our result on optimality ([Theorem 1](#)) holds for far more generic contact processes. In particular, mobile users need not contact each other in an independent manner, and the inter-contact times need not be exponential.

Formally assessing the effect of social network properties such as node centrality and the strength of relationships on the performance DTN systems remains largely an open question. Though our focus is on a different application than [7] and [8], our work rigorously relates the social network's properties to the system's scalability and to algorithms for finding the optimal allocation. As such, it strengthens the foundations of research in the area of mobile social networks.

## Results & Contributions.

Our first contribution is showing that the service provider can allocate its downlink capacity in an optimal manner. More precisely, the provider can compute a downlink rate allocation that maximizes the network's social welfare -the aggregate utility over all users. We prove this in [1] by showing that the corresponding optimization problem is convex and, therefore, can be solved efficiently by gradient descent. Moreover, we give both a centralized and a distributed algorithm for computing the gradient; these can be used by the service provider to compute the optimal allocation, as we illustrate through an empirical study.

Our second contribution is proving that the system is scalable under the condition that the social network formed by the subscribers has a bounded edge expansion. In particular, even if the provider has a fixed downlink capacity, the age of the content available to users will grow slowly (as  $\log n$ ) as the number of users  $n$  increases. Our second result therefore identifies edge expansion as a key property of the social network that affects the system's scalability. Most importantly, it also implies that the provider can exploit this social network to offer the service in spite of using only limited resources, while sustaining a slow quality degradation under system growth.

Our empirical study uses two real-life mobility traces, spanning over different time scales (a few hours and several days, respectively). We compute the optimal downlink allocation and compare it to several simple heuristics, illustrating its dependence on system parameters. An interesting outcome of our study is that the intuition that the “most social” users should receive more frequent updates can in fact be wrong: under certain conditions, it is actually optimal to allocate no bandwidth to the most social users in the system.

## Model

Before presenting our results, we first give a brief description of our model. We assume that the service provider generates new content updates according to a Poisson process with rate  $\mu$ . Moreover, assume that there are  $n$  mobile users and that each update is injected by the service provider to

only one user. As a result, each user receives updates according to a Poisson process with rate  $x_i$ , such that 
$$\sum_{i=1}^n x_i \leq \mu.$$

Let  $Y_i$  be the age of the content at user  $i$ . Moreover, assume that each user  $i$  receives a utility  $u_i(Y_i)$  from the content, which only depends on the age  $Y_i$ . As a user older content should be less valuable, we only consider non-increasing utilities  $u_i$ . Under these assumptions, a natural goal for the service provider is to maximize the aggregate utility over all users, i.e., the social welfare:

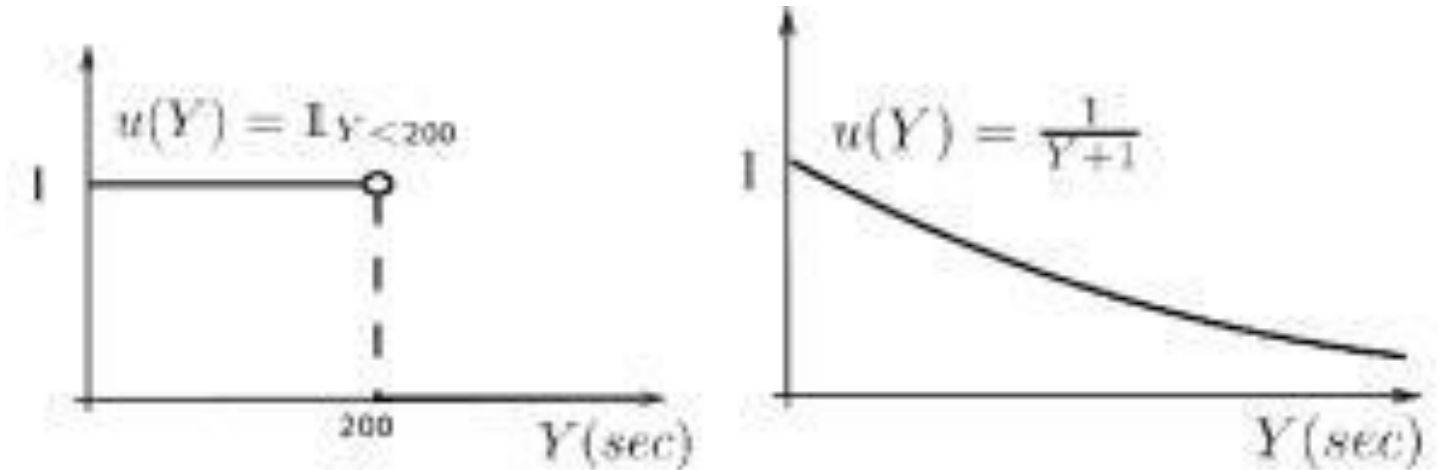


Illustration 2: User utilities in our experiments.

### SOCIAL WELFARE MAXIMIZATION

$$f(\vec{x}) = \sum_{i=1}^n E_{\vec{x}} u_i(Y_i) \quad (*) \quad \text{Maximize:}$$

$$\text{Subject to:} \quad \sum_{i=1}^n x_i \leq \mu, \quad x_i \geq 0, \quad 1 \leq i \leq n.$$

Note that the optimal allocation will depend on  $\mu$ , the utilities  $u_i$ , as well as the process describing the opportunistic contacts between users. The latter is governed by the users' social behavior, as this behavior determines who meets whom, and when.

### Main Results

Our first main result concerns the solution of the SOCIAL WELFARE MAXIMIZATION problem. We prove the following theorem, whose generality is surprising: obtaining an optimal allocation is feasible for (a) all non-increasing utilities and (b) for all stationary ergodic contact processes between users. The joint contact process exhibits this property if, for example, the marginal contact processes between distinct pairs of users are independent renewal processes (not necessarily Poisson). Independence however is not strictly necessary. For example, users can move according to random waypoint model independently or in groups. In either case, the joint contact process will be stationary ergodic, although the marginal contact processes for distinct pairs will not be independent.

**Theorem 1.** Assume that the process describing contacts between users is stationary ergodic. Then, SOCIAL WELFARE MAXIMIZATION is a convex optimization problem. In particular, the objective function  $f$ , given by (\*), is concave.

Theorem 1 implies that the optimization problem can be solved by gradient descent. In [1], we outline how gradient descent can be implemented by the service provider in a centralized fashion. Our method requires that the service provider knows the user utilities and that the users maintain contact logs of other users they encounter, which they periodically report to the service provider.

It is important to note however that, in fact, the optimal allocation can be computed without knowing the user utilities or collecting contact logs. In [1], we also show that the users themselves can estimate the gradient of the objective function in a fully distributed manner. In brief, our distributed scheme works by making users generate “dummy” update messages, that are propagated along with the real updates. To estimate the effect that the rate  $x_j$  has on the expected utility  $E[u_i(Y_i)]$  of user  $i$ , the user can simply monitor the effect that these dummy updates have on its utility.

Our second main result addresses system scalability. For this result we assume, as in [5] and [6], that the contact processes among pairs of users  $i, j=1, \dots, n$ , are independent Poisson processes with rates  $q_{ij} \geq 0$ . In this case, our system can be represented by a complete, weighted, undirected graph  $G(V, E)$  whose vertex set is  $V = \{1, \dots, n\}$  and whose edge weights are  $q_{ij}$ ,  $i, j \in V$ . Our result, summarized in the following theorem, relates the scalability of our system to the edge expansion  $h_G$  [9] of the graph  $G$ , defined as:

$$h_G = \min_{A \subseteq V} \frac{\sum_{i \in A, j \in A^c} q_{ij}}{\min(|A|, |A^c|)}.$$

Theorem 2. Assume that the processes describing the contacts between any pair of users  $i$  and  $j$ , where  $i, j=1, \dots, n$ , are independent Poisson processes.

the rate allocation is the uniform allocation (i.e.,  $x_i = \frac{\mu}{n}$  If  $i=1, \dots, n$ ), then the expected age seen by any user  $i$  satisfies where  $h_G$  the edge

$$E_{\bar{x}}[Y_i] \leq \frac{2}{\mu} (2e^{-1/2} + \log(n)) + h_G^{-1} \log(n) \quad \text{expansion of } G.$$

Theorem 2 indicates that our system scales very well if the edge expansion  $h_G$  is bounded away from zero, as the system size  $n$  grows. In such a system, even though the service provider commits only a constant amount of bandwidth ( $\mu$ ) to serve its users, sharing updates guarantees an age that grows slowly with  $n$  (as  $\log n$ ). Graphs having the constant expansion property are called expander graphs [9]. As we discuss below, in [2], we investigate the applicability of the above theorem on real data traces.

## Empirical Study of Optimal Allocation

In [1], we used the centralized algorithm to find the optimal rate allocation for two human mobility traces collected by other researchers. The Infocom06 data set contains Bluetooth contacts between iMotes distributed to participants of the three-day Infocom '06 conference. We focused on a 10 hour period during the first day of the conference. In the MIT data set, collected by the Reality-Mining project cite{reality-mining}, participants carried GSM enabled cellphones over a period of 9 months. We assume that two phones are in contact when they share the same GSM base station. We exclude 12 users from our analysis, as they were isolated. Moreover, due to memory constraints, we limit our analysis of the MIT dataset to an 80 day period. For simplicity, we assumed that every user has the same utility  $u_i(Y_i) = u(Y_i)$ , where  $u$  is one of the functions shown in [Illustration 2](#).

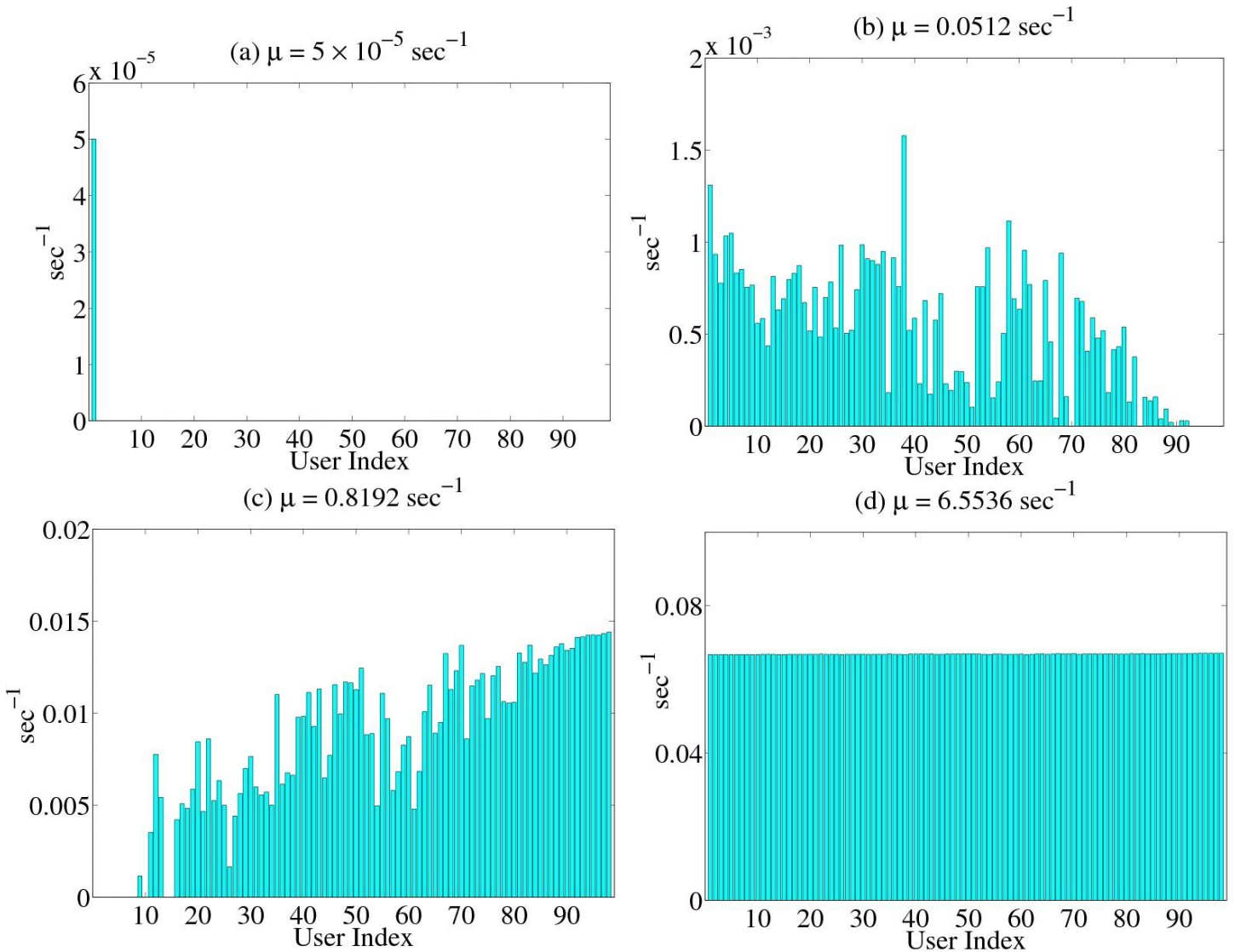


Illustration 3: The optimal allocation for the step utility function in the Infocom06 dataset. Users are indexed according to their contact rates (i.e., how social they are) in decreasing order.

[Illustration 3](#) presents the optimal rate allocation obtained with the step function utility in the Infocom06 data set, for different values of  $\mu$ . As shown in Illustrations 3(a) and 3(b), for small  $\mu$ , the optimal allocation tends to be biased towards the “most social” users in the system, i.e., the ones with the highest aggregate contact rates. In Infocom06, for both utility functions, whenever  $\mu$  is less than  $6 \times 10^{-3}$  updates per second, the optimal allocation concentrates on the “most central” user, from which content is disseminated to all users the fastest. In Infocom06, this node also has the highest aggregate contact rate -in other words, it is also the “most social” user. The optimal allocation also concentrates on a single “most central” user for MIT, when  $\mu$  is less than  $4 \times 10^{-4} \text{sec}^{-1}$ , though this user does not have the highest contact rate in this data set. This indicates that the “most social” user need not also be the “most central”.

For  $\mu$  very large, the optimal rate allocation becomes uniform. The improvement provided by content sharing becomes negligible and the system behaves as if nodes were isolated and no sharing takes place. It can be verified that, by [Theorem 1](#), if all users are isolated, the uniform allocation is optimal.

We observe an interesting phenomenon when  $\mu$  is between 0.2 and 0.8192  $\text{sec}^{-1}$  in the Infocom06 data set. In this region, the rate at the 8 “most social” users becomes zero. This indicates that the behavior of the “most central” user is fully reversed: while for low values of  $\mu$  this user acts as a global source of all incoming information for all users, in this region of  $\mu$  it receives no updates from the service provider. In other words, it is sometimes optimal to not push any updates to the most social users in the system but, instead, let them get the content indirectly!

[Illustration 4](#) shows the ratio between the social welfare achieved under three heuristics and the optimal. We consider (a) the uniform allocation, (b) a skewed allocation, in which all the injection rate is concentrated at the “most social” user, and (c) an allocation in which each user receives an injection rate that is proportional to its aggregate contact rate.

The comparison of these allocations confirms our earlier observations. The skewed allocation performs well for small values of  $\mu$ , but not always optimally as it may not select the “most central” user. Uniform is always optimal for large values of  $\mu$ . Proportional is the best among the three for intermediate values of  $\mu$ .

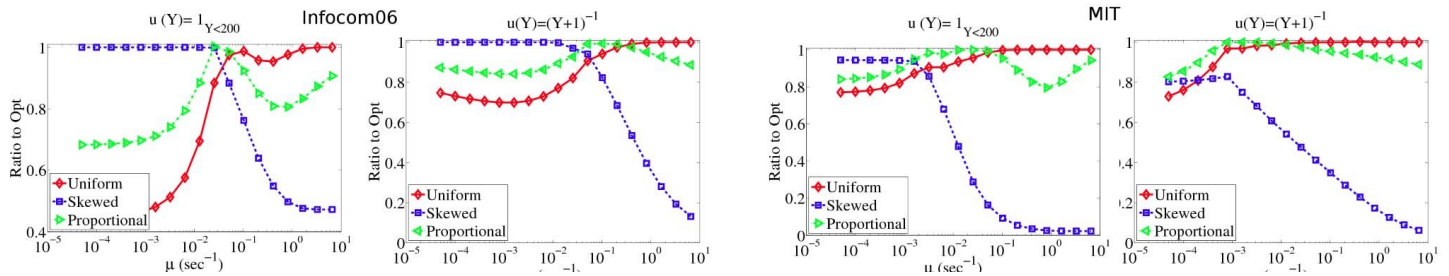


Illustration 4: Ratio of the social welfare under different heuristic allocations to the one achieved by the optimal.

In conclusion, our results highlight a transition depending on  $\mu$  for the optimal rate allocation from concentrated to uniform. The topology of the users' social network plays an important role (e.g., in selecting the “most central” nodes) which becomes more prominent when  $\mu$  takes intermediate values.

## Empirical Study of Edge Expansion

To illustrate the relevance of [Theorem 2](#), we investigated in [\[2\]](#) how foregoing certain opportunities for sharing content would affect our system. In particular, although the actual contact rate between two users  $i$  and  $j$  may be  $q_{ij}$ , a user may choose to utilize only a fraction  $q_{ij}' \leq q_{ij}$  for sharing updates. E.g., users may reduce their bandwidth consumption according to the following strategies:

- **Proportional(p)**. Every contact rate is scaled by  $0 \leq p \leq 1$ , i.e.,  $q_{ij}' = p q_{ij}$ . This can be implemented by ignoring contact events with probability  $1-p$  and using them for content sharing with probability  $p$ .
- **Keep Highest(k)**. Each user  $i$  restricts its contacts to only  $k$  users. These  $k$  users are the ones with the highest contact rates with this node.
- **Proportional Keep Highest(p)**. Each user  $i$  reduces its aggregate contact rate to  $p q_i$ , where  $0 \leq p \leq 1$ , by removing a rate  $(1-p)q_i$  in total from its neighbors with the lowest rate.
- **Proportional Keep Lowest(p)**. Each user  $i$  reduces its aggregate contact rate to  $p q_i$ , where  $0 \leq p \leq 1$ , by removing a rate  $(1-p)q_i$  in total from its neighbors with the highest rate.
-

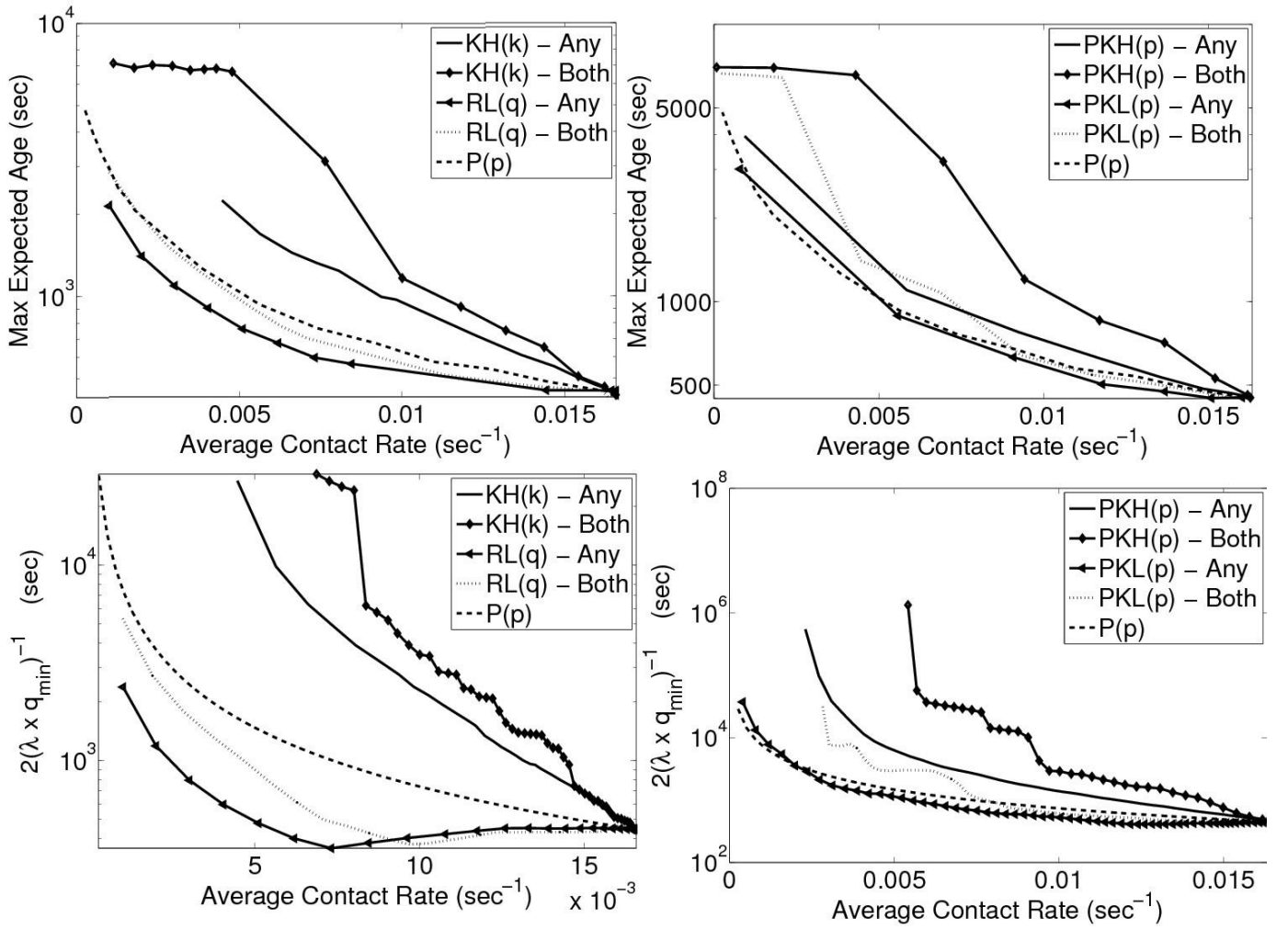


Illustration 5: Maximum age and inverse of expansion for different rate reduction strategies, in the Infocom06 dataset. The inverse of the expansion  $h_G^{-1}$  is approximated by  $2(\lambda q_{\min})^{-1}$ , where  $\lambda$  the spectral gap of the graph's Laplacian and  $q_{\min}$  is the minimum contact rate of all users.

**Rate Limit (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  $q_i' = q$ .

Except for Proportional(p), all strategies may yield asymmetric contacts between two users. We therefore consider two variants of each strategy. In the first, a contact between  $i$  and  $j$  is used if both users choose to use it. In the second, a contact is used if any of the two users chooses to use it.

In [2], we investigated how the maximum expected age of each user is affected by several strategies, including the five described above, for both the Infocom06 and the MIT datasets. We also investigated what effect these strategies have on edge expansion. Computing the edge expansion of a given graph is an NP-hard problem, so we approximated  $h_G$  using the bound  $h_G \geq 0.5 \lambda q_{\min}$ , where  $\lambda$  is the spectral gap of the graph's Laplacian and  $q_{\min}$  is the minimum contact rate of all users [9].

The top row of Illustration 5 shows the average content for different strategies in Infocom06. All results agree that the best performance trade-off is achieved by strategies that prioritize weak ties, i.e., connections with users that one does not meet very often. In particular, strategies may be sorted in decreasing order of efficiency as Rate-Limit, Proportional-Keep-Lowest, Proportional, Proportional-Keep-Highest and Keep-Highest. With the Rate Limit strategy, that achieves the best trade-off in both datasets, 75% of the contacts can be removed without even doubling the maximum content age.

The second row of Illustration shows the upper bound of  $h_G^{-1}$  (in terms of  $\lambda$  and  $q_{\min}$ ), 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 age for both data sets. We note that this is quite surprising, given that Theorem 2 merely establishes an upper bound and that our approximation overestimates  $h_G^{-1}$ . Moreover, it corroborates the result of Theorem 2 by indicating that, for these data sets, the maximum age is inherently linked to the edge expansion of  $G$ .

## Conclusions

Our results show that (a) content updates can be distributed over a mobile social network in a scalable way (b) the service provider can successfully exploit the social network to obtain an optimal allocation of its aggregate injection rate. Possible extensions of this work include the introduction of incentives to share as well as full decentralization, whereby no service provider exists and content is generated by the users themselves.

## References

- [1] S. Ioannidis, A. Chaintreau, and Laurent Massoulié, “Optimal and Scalable Distribution of Content Updates over a Mobile Social Network”. In IEEE INFOCOM ,2009.
- [2] S. Ioannidis, A. Chaintreau. “On the Strength of Weak Ties in Mobile Social Networks”. In SNS, 2009.
- [3] D. Kempe, J. Kleinberg, and Éva Tardos, “Maximizing the Spread of Influence Through a Social Network”, in ACM KDD, 2003.
- [4] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, “Cost-Effective Outbreak Detection in Networks”, in ACM KDD, 2007.
- [5] E. Altman, G. Neglia, F. De Pellegrini, and D. Miorandi, “Decentralized Stochastic Control of Delay Tolerant Networks”. In IEEE INFOCOM, 2009. [6] W. Gao, Q. Li, B. Zhao, and G. Cao, “Multicasting in Delay Tolerant Networks: A Social Network Perspective”. In ACM MOBIHOC, 2009.
- [7] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, “A Socio-aware Overlay for Publish/Subscribe Communication in Delay Tolerant Networks”, in ACM MSWiM, 2007.
- [8] 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 UbiComp, 2007.
- [9] F.R.K. Chung, “Spectral Graph Theory”. American Mathematical Society, 1997.