

Analysing Delay-tolerant Networks with Correlated Mobility

Mikael Asplund, Simin Nadjm-Tehrani

Department of Computer and Information Science
Linköping University
`{mikael.asplund,simin-nadjm.tehrani}@liu.se`

Abstract. Given a mobility pattern that entails intermittent wireless ad hoc connectivity, what is the best message delivery ratio and latency that can be achieved for a delay-tolerant routing protocol? We address this question by introducing a general scheme for deriving the routing latency distribution for a given mobility trace. Prior work on determining latency distributions has focused on models where the node mobility is characterised by independent contacts between nodes. We demonstrate through simulations with synthetic and real data traces that such models fail to predict the routing latency for cases with heterogeneous and correlated mobility. We demonstrate that our approach, which is based on characterising mobility through a colouring process, achieves a very good fit to simulated results also for such complex mobility patterns.

Keywords: Latency, Delay-tolerant networks, Correlated Mobility, Connectivity

1 Introduction

Delay- and disruption-tolerant networks represent an extreme end of systems in which a connected network cannot be relied upon. Instead, messages are propagated using a store-carry-forward mechanism. Such networks can have applications for disaster area management [4], vehicular networks [19], and environmental monitoring [17]. These systems offer many challenges and have been extensively studied by the research community [1, 22, 23, 26].

Recent results indicate that to the extent that delay-tolerant networks will be found on a larger scale, they will definitely be composed of islands of connectivity, that is, some parts that are well-connected and some parts that are sparse. This in turn implies correlated contact patterns [2, 11]. Most existing analytical delay performance models fail to capture such scenarios, since they assume independent node contacts. Moreover, although there are analyses done also for quite complex mobility models [8, 9], it is not obvious how one should go about to map such models from real traces.

We extend previous results by studying the routing latency distribution for heterogeneous mobility movements. Our analytical model incorporates a colouring technique for information propagation to derive the latency distribution for

an epidemic routing algorithm for a quite general case. The key strength of our approach compared to other models of heterogeneous mobility is that we are able to extract the relevant data from a real trace and produce the routing latency distribution (not just expected latency). The results are verified with a simulation-based study where we consider both synthetic and real-life mobility traces. We show that while a model that assumes independent inter-contact times works well for simple synthetic models such as random waypoint it is not able to predict the routing performance for a heterogeneous mobility model whereas our analytical results match very well.

There are two main contributions in this paper. First, a scheme for deriving the routing latency distribution for complex heterogeneous mobility models and, second, an experimental evaluation and validation of our model and a comparison with a model that assumes homogeneous and independent mobility is presented. The key insight of the evaluation is that heterogeneous mobility can result in such a high correlation of contacts that theoretical results based on independent inter-contact times are no longer valid.

The rest of the paper is organised as follows. Section 2 describes the system model and the basic assumptions we make. Section 3 describes how to derive the routing latency distribution given knowledge of the colouring rate distribution. This latter distribution is discussed in Section 4, and we explain how it can be determined from mobility traces. Section 5 contains the experimental evaluation. Finally, Section 6 gives an overview of the related work and Section 7 concludes the paper.

2 System Model

Consider a system composed of N mobile nodes (some possibly stationary). Nodes can communicate when they are in contact¹ with each other. During the contact both nodes can send and receive messages. We focus on connection patterns and ignore effects of queueing and contentions. Moreover, since we are interested in intermittently connected networks, the time taken to transmit a message is assumed negligible in relation to the time taken to wait for new contacts. We call this assumption A.

We characterise the pattern with which contacts occur using a simple colouring process (similar to [22, 23]). Note that the colouring does not necessarily correspond to message dissemination, and should be seen only as an indication of node contact patterns. The basic idea is that if node A is coloured and subsequently comes in contact with node B, then node B will also become coloured (if not already coloured). The only restriction we make on the contact pattern (and thereby on the mobility of the nodes) is that the *incremental colouring times* should be independent. More specifically, given a colouring process that has coloured i nodes, the time to colour one more node is independent from the time taken to colour the earlier i nodes. We call this assumption B. Note that

¹ A contact is defined by a start and an end time between which two nodes are within communication range.

Table 1: Notation

N	Number of nodes in the system	$P(X)$	Probability of X being true
T_i	Random variable, the time taken for a randomly chosen colouring process to colour i nodes	Δ_i	Random variable, the time taken for a randomly chosen colouring process to colour one more node given i coloured nodes
R	Random variable, the message delivery time	$f_i(t)$	PDF of the random variable T_i
$f_{\Delta_i}(t)$	PDF of the random variable Δ_i	$F_i(t)$	CDF of the random variable T_i
$F_{\Delta_i}(t)$	CDF of the random variable Δ_i	$F_R(t)$	CDF of the random variable R

this is a much weaker restriction on the set of allowed mobility models compared to assuming independent inter-contact times.

We use a number of random variables to describe the colouring and routing processes, Table 1 summarises the most important notation. PDF is an abbreviation for probability density function and CDF stands for cumulative density function, these abbreviations are used throughout the paper.

Our analysis builds on ideal epidemic routing since it corresponds to the optimal performance any routing algorithm can achieve. Thus, these results provide a useful theoretical reference measure on what is good performance for a given mobility model. Such a reference can also be of practical use to decide whether the measured performance in some network is due to the network characteristics or to the protocol implementation. Moreover, this scheme can be extended to other routing protocols, for example using the techniques described by Resta and Santi [22].

3 Routing Latency

We now proceed to characterise the routing latency for epidemic routing in intermittently connected networks. We begin by determining the colouring time distribution which is then used to express the routing latency distribution.

3.1 Colouring Time

A colouring process (t_0, s) is characterised by a start time t_0 and a source node s from which the colouring process begins (thus, s becomes coloured at time t_0). Every time a coloured node comes in contact with an uncoloured node, the uncoloured node becomes coloured. Let T_i denote the random variable representing the time taken for a randomly chosen colouring process to colour i nodes.

Moreover, we let Δ_i denote the random variable that describes the time taken for a randomly chosen colouring process to colour *one more node* given that i nodes are already coloured. This means that we can express the time taken for a colouring process to reach $i + 1$ nodes as $T_{i+1} = T_i + \Delta_i$.

Note that since we start the process with one coloured node, the time to colour the first node is $T_1 = 0$, and the time to colour the second node is $T_2 = \Delta_1$. Slightly abbreviating standard notation we let $f_i(t)$ denote the PDF of the random variable T_i and let $f_{\Delta i}(t)$ be the PDF of Δ_i . For the purpose of this presentation we assume that the latter of these functions is given since it depends on the mobility of the nodes in the system. In Section 4 we show how to extract $f_{\Delta i}(t)$ from an existing contact trace. Assumption B from Section 2 states that T_i and Δ_i are independent, so the PDF of their sum can be expressed as the convolution of the PDFs of the respective variables [10]:

$$f_{i+1}(t) = (f_i * f_{\Delta i})(t) \quad (1)$$

Since we know the characteristic of $f_2(t)$ we can use equation (1) to iteratively calculate $f_3(t), f_4(t), f_5(t)$ and so on. The CDF for the variable T_i , here denoted by $F_i(t)$, can be computed in the standard manner from the PDF by integrating over all time points. Thus, assuming that colouring times are independent, it is straightforward to express the colouring time distribution $F_i(t)$ given knowledge of the PDF $f_{\Delta i}(t)$. In the next subsection, we show how to derive the routing latency distribution from $F_i(t)$.

3.2 Routing Latency and Delivery Ratio

Our aim now is to find the latency distribution for an ideal routing algorithm. So, consider a randomly chosen time t_0 , source node s and destination node $d \neq s$. Let R be the random variable that models the time to route a message from s to d using ideal epidemic routing. We will try to find the CDF of R , $F_R(t) = P(R \leq t)$. Clearly, given assumption A (i.e., that the queueing and transmission times can be neglected), this probability is the same as for d being one of the coloured nodes by the colouring process (t_0, s) after t time units.

Let C_t be the random variable that models the number of coloured nodes after t time units. If $C_t = i$ then the probability that d is coloured after t time units is $(i - 1)/(N - 1)$ since if we remove the source node s , there are $i - 1$ coloured nodes and $N - 1$ nodes in total. Thus, we can express $F_R(t)$ as:

$$F_R(t) = P(R \leq t) = \sum_{i=1}^N P(C_t = i) \cdot \frac{i - 1}{N - 1} \quad (2)$$

Now let's consider the probability $P(C_t = i)$ that the number of coloured nodes at time t equals i . This is the same as the probability that the time taken to inform i nodes is less than or equal to t minus the probability that $i + 1$ nodes can be reached in this time:

$$P(C_t = i) = P(T_i \leq t) - P(T_{i+1} \leq t) \quad (3)$$

Combining equations (2) and (3), and rewriting gives:

$$F_R(t) = \frac{1}{N - 1} \sum_{i=2}^N F_i(t) \quad (4)$$

Listing 1 GetRoutingLatencyDistribution

Input: f_{Δ_i} : Vector representing the PDF of Δ_i

```
1   $f_2 \leftarrow f_{\Delta_1}$ 
2  for  $i = 3 \dots N$ 
3     $f_i \leftarrow \text{CONV}(f_{i-1}, f_{\Delta_{i-1}})$  /* equation (1) */
4     $F_i \leftarrow \text{CUMSUM}(f_i)$ 
5   $F_R \leftarrow \frac{1}{N-1} \sum_{i=2}^N F_i$  /* equation (4) */
6  return  $F_R$ 
```

In summary, if we know the probability PDFs of the random variables Δ_i , we can use equation (1) to determine $f_i(t)$. Equation (4) will then give us the cumulative distribution function for the epidemic routing latency. Listing 1 shows an algorithmic representation of how to derive the distribution for R using discrete distributions. The procedure CONV and CUMSUM are standard Matlab functions and compute the convolution between two vectors and cumulative vector sum respectively.

By knowing R we can easily deduce the delivery ratio of a protocol given a certain time-to-live (TTL) for each packet. The probability that a message with TTL of T will reach its destination is simply $F_R(T)$ (i.e., the probability that the message will be delivered within time T).

4 Colouring Rate

Having derived the routing latency distribution based on knowledge of the distribution of the incremental colouring time Δ_i we now proceed to show how to find this latter distribution.

We consider two cases, when the mobility is homogeneous, and the more interesting heterogeneous case. By homogeneous we mean that the pairwise inter-contact times (i.e., the time between contacts) are identical and independently distributed (often abbreviated iid). The homogeneous case is not really novel in this context and is provided here briefly in order to explain the baselines we have used and to show that this case is also covered by our general approach.

4.1 Homogeneous Mobility

For the particular case of homogeneous mobility we make three additional assumptions commonly used to analyse homogeneous mobility [7, 12]. (H1) The duration of contacts is negligible compared to the waiting times, (H2) the inter-contact time has a finite expectation, and (H3) pair-wise contacts are independent.

Now consider a set of coloured nodes that wait for a new contact to appear so that a new node can become coloured. The time they have to wait is the smallest of all pairwise waiting times for all pairs where one node in the pair

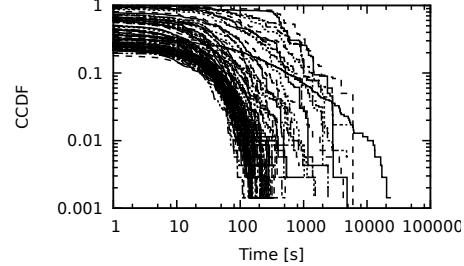


Fig. 1: Complementary Cumulative Distribution Functions (CCDF) of Δ_i

is coloured and one node is uncoloured. If i nodes are coloured, then there are $i(N - i)$ such pairs. Given assumption H3, we can express the CDF of Δ_i as:

$$F_{\Delta i}(t) = P(\Delta_i \leq t) = 1 - (1 - F_\tau(t))^{i(N-i)} \quad (5)$$

where $F_\tau(t)$ is the cumulative distribution of the residual² inter-contact time between two nodes. We refer to Karagiannis et al. [12] for further explanation and how to derive the residual distribution from the inter-contact distribution. If the inter-contact time is exponentially distributed with rate λ , then the residual waiting time is also exponentially distributed with the same rate and the incremental colouring time Δ_i will be exponentially distributed with rate $\lambda i(N - i)$.

4.2 Heterogeneous Mobility

If node contacts are not independent, then deriving an expression for the colouring distribution Δ_i will be more challenging. We now proceed to present a first simple model for approximating it from real heterogeneous traces.

In order to explain the rationale behind the model we first show some data from a real-life trace based on the movement of taxis in the San Francisco area. The trace was collected by Piorkowski et al. [21] based on data made available by the cabspotting project during May 2008 and we used a subset of the first 100 vehicles from the trace. In the simulation each taxi was assumed to have a wifi device with a range of 550m.

Fig. 1 shows the Complementary Cumulative Distribution Function (CCDF) of each Δ_i (recall that i corresponds to the number of already coloured nodes) for the San Francisco cab scenario. We obtained this data by running 700 of colouring processes on the contact trace and logging the time taken to colour the next node. The plot uses a logarithmic scale on both axes to highlight the characteristics of the distribution. This shows that they exhibit an exponential decay (i.e., it approaches 0 fast, indicated by the sharp drop of the curves.).

² The residual inter-contact time refers to the time left to the next contact from a randomly chosen time t , as opposed to the time to the next contact measured from the previous contact time.

The second phenomena that we have observed is that due to clustering of nodes, it is often the case that the next node can be coloured without any waiting time at all. Based on these two basic principles we conjecture that the colouring time can be modelled as either being zero with a certain probability, or with a waiting time that is exponentially distributed.

If i nodes have been coloured, then we let $Con(i)$ denote the probability that one of those i nodes is connected to an uncoloured node (thereby allowing an immediate colouring of the next node). Further we let $f_{\text{Exp}}(t, \lambda_i)$ denote the PDF of the exponential distribution with rate λ_i . Then, we let the PDF of the simple colouring distribution model be expressed as:

$$f_{\Delta i} = \begin{cases} Con(i) & \text{if } t = 0 \\ (1 - Con(i)) * f_{\text{Exp}}(t, \lambda_i) & \text{otherwise} \end{cases} \quad (6)$$

While this is clearly a simple model, it can be seen as a first step towards modelling the colouring distribution and seems to work well enough for the scenarios we have studied. We believe that further work is needed to better understand how the colouring distribution is affected by different mobility conditions. Note also that our general scheme is not tied to this particular model and allows further refinements.

5 Evaluation

To validate our model and to test whether it actually provides any added value compared to existing models we performed a series of simulation-based experiments. We used three different mobility models, the random waypoint mobility model, a model based on a map of Helsinki and a real-world trace from the cabs in the San Francisco area. After explaining the experiment setup we give the details and results for each of these models. Finally, we relate our findings on the effects of heterogeneity for these cases.

We used the ONE Simulator [14] to empirically find the ideal epidemic routing latency distribution for the three different mobility models. For each mobility model we ran the simulation 50 times. For the first 40000 seconds a new message with random source and destination was sent every 50 to 100 seconds. The simulation length was sufficiently long for all messages to be delivered. We used small messages of size 1 byte, and channel bandwidth of 10Mb/s.

In addition to the simulated results we used two different theoretical models to predict the latency distribution:

Colouring Rate: This model uses equation (6) from Section 4.2 to model the colouring times. The necessary parameters $Con(i)$ and λ_i are estimated from the trace file by sampling.

Homogeneous: This model assumes independent and exponentially³ distributed inter-contact times which are used to compute $f_{\Delta i}$ as described in Sec-

³ We also obtained nearly identical results when estimating the inter-contact distribution from the mobility trace, which we have excluded for lack of space.

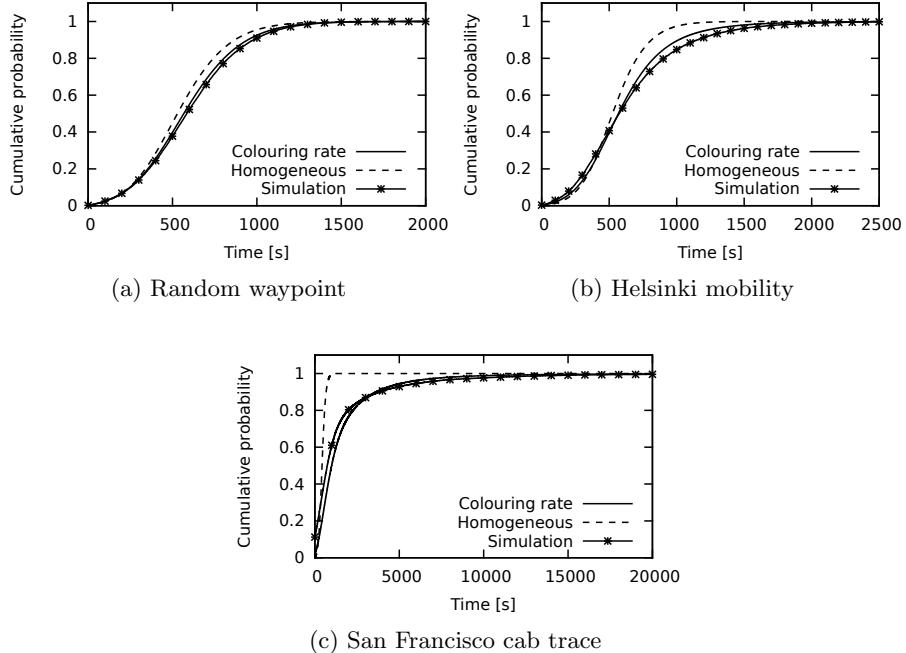


Fig. 2: Routing latency

tion 4.1. This has been a popular model for analysing properties of delay-tolerant networks [22, 23, 26].

In order not to get a biased value for the inter-contact time distributions due to a too short sampling period, we analysed contacts from 200 000 seconds of simulation. To further reduce the effect of bias we use Kaplan-Meier estimation as suggested by Zhang et al. [26].

5.1 Effect of Mobility

Random Waypoint Mobility. In order to validate our model against already known results, we start with considering the random waypoint mobility model. Despite its many weaknesses [2, 25], this model of mobility is still very popular model for evaluating ad hoc communication protocols and frameworks. The network was composed of 60 nodes moving in an area of $5km \times 5km$, each having a wireless range of $100m$. The speed of nodes was constant $10m/s$ with no pause time.

Fig. 2a shows the results of the two theoretical models and the simulation. The graph shows the cumulative probability distribution (i.e., the probability that a message will have been delivered within the time given on the x axis).

As expected, both models manage to predict the simulated results fairly well. In fact, the exponential nature of the inter-contact times of RWP is well understood and since the heterogeneous model is more general, we were expecting similar results.

Helsinki Mobility. We now turn to a more realistic and interesting mobility model, the Helsinki mobility model as introduced by Keränen and Ott [13]. The model is based on movements in the Helsinki downtown area. The 126 nodes is a mix of pedestrians, cars, and trams, and the move in the downtown Helsinki area (4500x3400 m). We used a transmission range of 50 meters for all devices. Fig. 2b shows the results. Again both theoretical models achieve reasonable results. However, due to the partly heterogeneous nature of the mobility model, the homogeneous model differs somewhat more from the simulated result. In particular, we see that the s-shape is more sharp compared to the observed data. We further discuss possible explanations for this in Section 5.2.

San Francisco Cabs. Finally, the last mobility trace we have analysed is a real-life trace based on the movement of taxis in the San Francisco area as explained in Section 4.2. Fig. 2c shows the results. In this case the homogeneous model fails to capture the routing latency that can be observed in simulation. However, the heterogeneous model based on equation (6) is still quite accurate. We were surprised to find such a big difference between the simulated data and the homogeneous model. Something is clearly very different in this trace compared to the synthetic mobility models. An estimate of the fraction of messages being delivered within an average latency of 2500s in such a scenario would be misleadingly optimistic by 20%.

5.2 The Effects of Heterogeneity

In the previous subsection we have seen that the accuracy of the homogeneous model is high for the random waypoint model, but is lower for the Helsinki model and completely fails for the San Francisco cab trace. In this subsection we present our investigation into why this is the case. We proceed by identifying four different aspects of how this model differs from reality.

Correlation. We begin with the most striking fact of the results presented so far. The homogeneous model is way off in predicting the routing latency distribution in the San Francisco case. There are a number of different ways that one can try to explain this, but we believe that the most important one has to do with correlation (i.e., non-independence) of events. The main assumption that makes equation (5) possible, and thereby the homogeneous model is that the contacts between different pairs of node are independent from each other. However, this seems to be a false assumption.

We analysed the contact patterns of the three different mobility models by considering the residual inter-contact times for each node during a period of 20000 seconds. Fig. 3a shows the percentage of nodes who's average correlation among its contacts is higher than a given value (i.e., it is the complementary CDF of nodes having a given average correlation). If the pairwise contacts are independent, they will have no (or very low) average correlation and we would

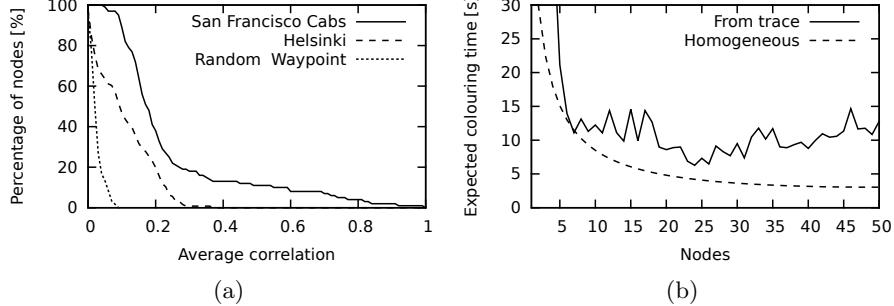


Fig. 3: (a) Correlation of contacts, (b) Time to colour one more node in the San Francisco trace

expect to see a sharp decay of the curve in the beginning of the graph. This is also what we see for the random waypoint model. Since the nodes move around completely independently from each other, the contacts also become independent. The Helsinki trace shows a higher degree of correlation, but not as significant as for the San Francisco cab case. In this case 40% of the nodes have an *average* correlation of their contacts which is higher than 0.2 (a correlation of 1 would mean that all contacts are completely synchronised). This shows a high degree of dependence and we believe provides an explanation of the result we have seen in Section 5.1.

Note that correlated mobility does not necessarily lead to slower message propagation, in fact there are results indicating the contrary [8]. What we have seen is that the *prediction* of the latency becomes too optimistic when not taking correlation into account. If the model assumes that contacts are “evenly” spread out over time, whereas in reality they come in clusters, the results of the model will not be accurate.

Lack of Expansion. The second prominent effect is what we choose to call lack of expansion (motivated by the close connection to expander graphs [3]). This means that the rate of the colouring process seems not to correspond to the number of coloured nodes. Fig. 3b shows the expected time to colour one more node for the San Francisco trace. The x-axis represents the number of nodes already coloured (up to half the number of nodes). We can see that the homogeneous model predicts that the time decreases (i.e., the rate of colouring increases) as the number of coloured nodes increase. On the other hand, the data based on sampling the distribution of Δ_i from the mobility trace file (indicated as “From trace” in the figure) shows that after the first 5-10 nodes have been coloured, the rate is more or less independent from the number of coloured nodes. We believe that this is partly due to the fact that most of the node mobility is relatively local and that nodes are often stationary for long periods of time.

Slow Finish. Another effect that can be observed is that in some rare cases it can take a very long time for a message to reach its destination. For example

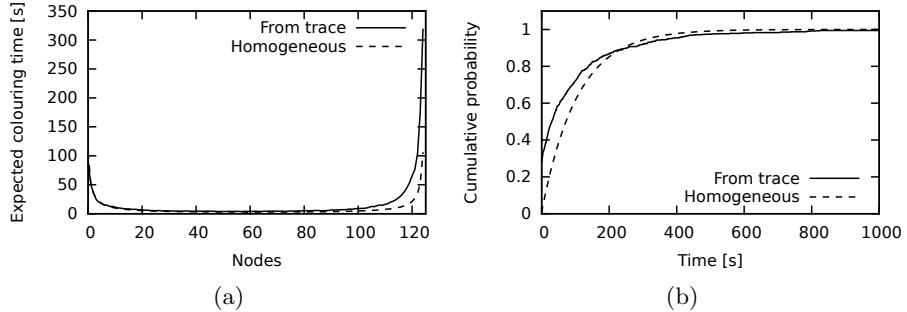


Fig. 4: (a) Time to colour one more node with the Helsinki mobility model, (b) CDF of the time to colour the second node

in Fig. 2c, even after 10000 seconds not all messages have been delivered to their destinations. This has to do with the fact that the time to colour all nodes take significantly longer time than to colour *almost* all the nodes. The models based on independent contacts predict that it takes the same amount of time to colour the second node as it takes to colour the last node. In both cases there are $N - 1$ possible node pairs that can meet and result in a colouring. However, we have seen that in reality colouring the last node takes significantly longer (on average). Fig. 4a shows the effect for the Helsinki trace, by plotting the expected colouring time as a function of the number of coloured nodes. While the homogeneous model is completely symmetrical around the middle, the actual data shows that it takes roughly three times longer to reach the last node than to reach the second node.

Fast Start. Finally, we consider why the homogeneous model predict a lower probability for delivering messages fast. This can be seen in both the Helsinki and San Francisco cases, but is more distinct in the former case. It can be seen visually in Fig. 2b in that the homogeneous model has a slightly flatter start compared to the other curves. This is because there is a chance that when a message is created, the node at which it is created has a number of neighbours. Thus, the message will not need to wait any time at all before being transmitted. Or if we express it as a colouring process, the time to colour the second node is sometimes zero. For a model based on inter-contact times, this is not considered.

Fig. 4b shows the CDF of T_2 , (i.e., the time taken to colour the second node) for the Helsinki case with the colouring rate and homogeneous model. We see that both curves are similar (the expected value for T_2 is the same for both models) but that the start value differs. That is, in the homogeneous model, it is predicted that the chance that the second node is immediately coloured is zero, whereas in fact it is roughly 0.3. Recall that the colouring time only reflects the contact patterns of the mobility and does not consider message transmission delays.

In this section we have seen how heterogeneous mobility causes correlated contacts and how that affects predictions of routing latency. Our model which is based on colouring rate of nodes was the only model able to accurately predict the routing latency distribution in these cases.

6 Related Works

There is a rich body of work discussing detailed analytical models for latency and delivery ratio in delay-tolerant networks. The work ranges from experimentally grounded papers aiming to find models and frameworks that fit to observed data to more abstract models dealing with asymptotic bounds on information propagation. Many of these approaches are based on or inspired by epidemiological models [15]. We have previously characterised the worst-case latency of broadcast for such networks using expander graph techniques [3].

Closest to our work in this paper is that of Resta and Santi [22], where the authors present an analytical framework for predicting routing performance in delay-tolerant networks. The authors analyse epidemic and two-hops routing using a colouring process under similar assumptions as in our paper. The main difference is that our work considers heterogeneous node mobility (including correlated inter-contact times), whereas the work by Resta and Santi assumes independent exponential inter-contact times.

Zhang et al. [26] analyse epidemic routing taking into account more factors such as limited buffer space and signalling. Their model is based on differential equations also assuming independent exponentially distributed inter-contact times. A similar technique is used by Altman et al. [1], and extended to deal with multiple classes of mobility movements by Spyropoulos et al. [24].

Kuiper and Nadjm-Tehrani [16] present a quite different approach for analysing performance of geographic routing. Their framework can be used based on abstract mobility and protocol models as well as extracting distributions for arbitrary mobility models and protocols from simulation data. The main application area for this model is geographic routing where waiting and forwarding are naturally the two modes of operation in routing.

The assumption of exponential inter-contact times was first challenged by Chaintreau et al. [7] who observed a power law of the distribution for a set of real mobility traces (i.e., meaning that there is a relatively high likelihood of very long inter-contact times). Later work by Karagiannis et al. [12] as well as Zhu et al. [27] showed that the power law applied only for a part of the distributions and that from a certain time point, the exponential model better explains the data. Pasarella and Conti [20] present a model suggesting that an aggregate power law distribution can in fact be the result of pairs with different but still independent exponentially distributed contacts. Such heterogeneous but still independent contact patterns have also been analysed in terms of delay performance by Lee and Eun [18].

Our work on the other hand, suggests that the exact characteristic of the inter-contact distribution is less relevant when contacts are not independent.

Correlated and heterogeneous mobility and the effect on routing have recently been discussed in several papers [6, 5, 8, 11], but to our knowledge, we are the first to provide a framework that accurately captures the routing latency distribution for real traces with heterogeneous and correlated movements.

7 Conclusions and Future Work

We have presented a mathematical model for determining the routing latency distribution in intermittently connected networks based on trace analysis. The basic idea that we have built upon is that the speed of a colouring process captures the dynamic connectivity of such networks. This was confirmed by a set of simulation-based experiments where we demonstrated that our model matched the simulation results very well. On the other hand, the models based on independent and homogeneous contacts did not provide accurate results except for the case with the random waypoint mobility model.

Our scheme allows accurate analysis of a much wider range of mobility models than previously possible. This analytical technique also has the possibility to increase our understanding of the connection between mobility and routing performance, potentially leading to new mobility metrics and classifications. We used a rough estimation-based model for the colouring distribution, and there is certainly room for considering other ways of expressing these distributions.

There are several possible extensions to this work. First, it would be interesting to study the accuracy of the analysis in the context of other routing paradigms such as social and geographic routing, as well as considering effects of limited bandwidth and buffers. Moreover, the effects of correlation of node contacts should be further investigated by analysing other real-life traces, also considering under which circumstances our assumption of independent colouring times is valid.

8 Acknowledgements

This work was supported by the Swedish Research Council (VR) grant 2008-4667. During the final stages of preparing the manuscript the first author was supported, in part, by Science Foundation Ireland grant 10/CE/I1855.

References

1. E. Altman, T. Basar, and F. D. Pellegrini. Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks. *Perform. Eval.*, 67(4), 2010. doi: 10.1016/j.peva.2009.09.001.
2. N. Aschenbruck, A. Munjal, and T. Camp. Trace-based mobility modeling for multi-hop wireless networks. *Comput. Commun.*, 34(6), 2010. doi: 10.1016/j.comcom.2010.11.002.
3. M. Asplund. *Disconnected Discoveries: Availability Studies in Partitioned Networks*. PhD thesis, Linköping University, 2010. [http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-60553](http://urn.kb.se/resolve?urn=urn:nbn:se:liu.diva-60553).
4. M. Asplund and S. Nadjm-Tehrani. A partition-tolerant manycast algorithm for disaster area networks. In *28th International Symposium on Reliable Distributed Systems (SRDS)*. IEEE, 2009. doi: 10.1109/SRDS.2009.16.

5. E. Bulut, S. Geyik, and B. Szymanski. Efficient routing in delay tolerant networks with correlated node mobility. In *Mobile Adhoc and Sensor Systems (MASS), 2010 IEEE 7th International Conference on*, 2010. doi: 10.1109/MASS.2010.5663962.
6. H. Cai and D. Y. Eun. Toward stochastic anatomy of inter-meeting time distribution under general mobility models. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '08. ACM, 2008. doi: 10.1145/1374618.1374655.
7. A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Trans. Mobile Comput.*, 6(6), 2007. doi: 10.1109/TMC.2007.1060.
8. D. Ciullo, V. Martina, M. Garetto, and E. Leonardi. Impact of correlated mobility on delay-throughput performance in mobile ad hoc networks. *Networking, IEEE/ACM Transactions on*, 19(6), 2011. doi: 10.1109/TNET.2011.2140128.
9. M. Garetto, P. Giaccone, and E. Leonardi. Capacity scaling in delay tolerant networks with heterogeneous mobile nodes. In *Proc. 8th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*. ACM, 2007. doi: 10.1145/1288107.1288114.
10. C. M. Grinstead and J. L. Snell. *Introduction to Probability*. American Mathematical Society, 1997.
11. T. Hossmann, T. Spyropoulos, and F. Legendre. Putting contacts into context: Mobility modeling beyond inter-contact times. In *Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 11)*. ACM, 2011.
12. T. Karagiannis, J.-Y. L. Boudec, and M. Vojnović. Power law and exponential decay of intercontact times between mobile devices. *IEEE Transactions on Mobile Computing*, 9, 2010. doi: 10.1109/TMC.2010.99.
13. A. Keränen and J. Ott. Increasing reality for dtn protocol simulations. Technical report, Helsinki University of Technology, Networking Laboratory, 2007.
14. A. Keränen, J. Ott, and T. Kärkkäinen. The ONE simulator for dtn protocol evaluation. In *Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, Simutools '09. ICST, 2009. doi: 10.4108/ICST.SIMUTOOLS2009.5674.
15. A. Khelil, C. Becker, J. Tian, and K. Rothermel. An epidemic model for information diffusion in manets. In *Proc. 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems (MSWiM)*. ACM, 2002. doi: 10.1145/570758.570768.
16. E. Kuiper, S. Nadjm-Tehrani, and D. Yuan. A framework for performance analysis of geographical delay-tolerant routing. *EURASIP Journal on Wireless Communications and Networking*, 2012. To appear.
17. S. Lahde, M. Doering, W.-B. Pöttner, G. Lammert, and L. Wolf. A practical analysis of communication characteristics for mobile and distributed pollution measurements on the road. *Wireless Communications and Mobile Computing*, 7(10), 2007. doi: 10.1002/wcm.522.
18. C.-H. Lee and D. Y. Eun. Exploiting heterogeneity in mobile opportunistic networks: An analytic approach. In *Sensor Mesh and Ad Hoc Communications and Networks (SECON), 2010 7th Annual IEEE Communications Society Conference on*, 2010. doi: 10.1109/SECON.2010.5508265.
19. R. Lu, X. Lin, and X. Shen. Spring: A social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks. In *INFOCOM, 2010 Proceedings IEEE*, 2010. doi: 10.1109/INFCOM.2010.5462161.
20. A. Passarella and M. Conti. Characterising aggregate inter-contact times in heterogeneous opportunistic networks. In J. Domingo-Pascual, P. Manzoni, S. Palazzo, A. Pont, and C. Scoglio, editors, *NETWORKING 2011*, volume 6641 of *Lecture Notes in Computer Science*, pages 301–313. Springer Berlin / Heidelberg, 2011. doi: 10.1007/978-3-642-20798-3_23.
21. M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser. A parsimonious model of mobile partitioned networks with clustering. In *First International Conference on Communication Systems and Networks (COMSNETS)*. IEEE, 2009. doi: 10.1109/COMSNETS.2009.4808865.
22. G. Resta and P. Santi. A framework for routing performance analysis in delay tolerant networks with application to noncooperative networks. *Parallel and Distributed Systems, IEEE Transactions on*, 23(1), 2012. doi: 10.1109/TPDS.2011.99.
23. T. Spyropoulos, K. Psounis, and C. Raghavendra. Efficient routing in intermittently connected mobile networks: The single-copy case. *IEEE/ACM Trans. Netw.*, 16(1), 2008. doi: 10.1109/TNET.2007.897962.
24. T. Spyropoulos, T. Turletti, and K. Obraczka. Routing in delay-tolerant networks comprising heterogeneous node populations. *IEEE Trans. Mobile Comput.*, 8(8), 2009. doi: 10.1109/TMC.2008.172.
25. J. Yoon, M. Liu, and B. Noble. Random waypoint considered harmful. In *Proc. INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*. IEEE, 2003. doi: 10.1109/INFCOM.2003.1208967.
26. X. Zhang, G. Neglia, J. Kurose, and D. Towsley. Performance modeling of epidemic routing. *Comput. Netw.*, 51(10), 2007. doi: 10.1016/j.comnet.2006.11.028.
27. H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L. Ni. Recognizing exponential inter-contact time in vanets. In *INFOCOM, 2010 Proceedings IEEE*, 2010. doi: 10.1109/INFCOM.2010.5462263.