Modeling Opportunistic Communication with Churn

Ljubica Pajevic*, Gunnar Karlsson

School of Electrical Engineering and Linnaeus Center ACCESS
KTH-Royal Institute of Technology
100 44 Stockholm, Sweden

Abstract

In opportunistic networking, characterizing contact patterns between mobile users is essential for assessing feasibility and performance of opportunistic applications. There has been significant efforts in deriving this characterization, based on observations and trace analyses; however, most of the previously established results were obtained by studying contact opportunities at large spatial and temporal scales. Moreover, the user population is considered to be constant: no user can join or leave the system. Yet, there are many examples of scenarios which do not fully adhere to the previous assumption and cannot be accurately described at large scales. Urban environments, such as smaller city districts, are characterized by highly dynamic user populations. We believe that scenarios with varying population require further investigation. In this paper, we present a novel modeling approach to study operation of opportunistic applications in scenarios where the population size is subjected to frequent changes, that is, it exhibits *churn*. We examine two location-based content sharing schemes: a purely opportunistic case and an infrastructure-supported content sharing scheme, for which we provide stochastic models based on stochastic differential equations (SDEs). We validate our models in five scenarios: a city area, subway station, conference, campus, and a scenario with a synthetic mobility model and we show that the models provide good representations of the investigated scenarios.

Keywords: Opportunistic networks, churn, stochastic models, stochastic differential equations, content sharing

1. Introduction

Opportunistic communication in urban areas is a promising solution for infrastructure-free location-based services and content sharing. The topic of urban opportunistic networking has resulted in a large body of research, with a majority of studies basing their results on scenarios where a fixed number of mobile users roams inside a closed area, following certain mobility patterns.

*Corresponding author
Email addresses: ljubica@kth.se (Ljubica Pajevic), gk@ee.kth.se (Gunnar Karlsson)

However, the size of the geographical space where the communication happens plays a significant role. As the region of interest shrinks—consider a small urban district, rather than an entire city—the assumption of a closed population will no longer hold: there will eventually be users arriving into the area, staying inside for a while and leaving. The population of users is no longer constant and such a system is known as a system with open population (*open system* in short), or in the networking parlance, a system with *churn*. Neglecting the effect of churn in specific scenarios can be hazardous.

Understanding contact patterns between humans and modeling these patterns realistically is essential for designing opportunistic communication schemes and evaluating their performance. While the research towards this aim has thrived over the past years, there has been little work considering opportunistic systems with churn. In such systems, user population is in constant change, thus previously obtained results (e.g. the distribution of inter-contact times) may not be reliable. This study takes a step towards increasing our understanding by characterizing user interactions in such dynamic systems.

Our initial work towards that aim has first been presented in [1]. Herein we extend our study by investigating an additional scenario and providing more detailed analysis. In particular, we examine the effects of node churn on opportunistic communication in the context of content sharing. By means of stochastic differential equations (SDE), we model and evaluate two schemes: the first scheme considers purely opportunistic communication and the second assumes some fixed infrastructure. These scenarios set the scene for the modelling approach which, we envision, can be adapted for a variety of other applications.

The second objective of this study is to gain insights on how user heterogeneity can be tackled in a more tractable way than has been done before. Departing from the traditional assumptions on the homogeneous contact patterns, the recent modeling attempts have resulted in notable contributions such as the works [2, 3]. In [2], the authors propose an analytic framework for evaluating performance of delay tolerant networks in terms of delivery delays, allowing different contact patterns for each pair of nodes in the system. The framework requires exact knowledge of the connectivity properties in a specific scenario and it relies on the exponential assumption for inter-contact times. More comprehensive is the performance modeling framework of Boldrini et al. [3], assuming general inter-contact times distributions. However, both frameworks are limited to closed populations.

The main contributions of this study are the following:

- We establish stochastic models for two epidemic content spreading schemes and we validate our models by means of simulation.
- We investigate five mobility scenarios including: realistic simulator-based traces, two sets of real-life traces and synthetic mobility. Then, we analyze the scenario-specific distributions of the contact rate and node sojourn time, and we compare these empirical distributions with modeling assumptions.

- We study contact patterns in systems with churn and provide an approximation of the contact rate for the system in the transient state.
- Based on the proposed modeling approach, we demonstrate an estimation method for the system parameters.

This paper is organized as follows. In Section 2 we describe the first content sharing scheme and present our modeling approach. Validation of this model is given in Section 3. The model for the second opportunistic application is developed in Section 4 and validated in Section 5. We review related work in Section 6 and summarize our most important findings in Section 7.

2. Opportunistic location-based content sharing: distributed case

First we consider an opportunistic service for sharing geographically localized contents. The service is targeted to mobile users in urban areas, where as an example of content type one can think of local news, tourist information, transportation schedules, traffic alerts, and the like. Content is geographically tagged to the region of relevance either explicitly e.g. given a set of geo-coordinates of a polygon that confines the region or the center and radius of the anchor zone [4], or implicitly by specifying geographic areas such as "central train station". Outside of the boundaries of the locale, content is considered irrelevant and will not be distributed.

The area of interest is characterized by frequent user arrivals and departures. We assume that mobile users are pedestrians equipped with mobile devices. Applications on user devices use services of a publish/subscribe middleware [5] to publish new contents, or to find peers within communication range and download or forward contents. Content can be classified into different distribution channels, however for the purpose of our study, we assume all users are subscribed to a single channel. Assume that initially there is a single user with a content item who wants to share it with other users in the area, without support of infrastructure. In addition to obtaining content, all nodes are willing to support further content spreading by contributing some amount of their resources for a limited time. This results in a fully distributed, purely opportunistic content sharing scheme. This scheme employs the *virtual storage effect* [6], and has also been referred to as *floating content* [7] or *hovering information* [8].

Users following certain mobility patterns will enter the area, move inside for a while and eventually leave. Since their movements are not bounded to the area, the feasibility and performance of such opportunistic scheme strongly depends on user mobility and contact patterns.

We are interested in answering questions such as:

1. Under what conditions the content is likely to persist in the area (that is, *survive*) for longer time, relying solely on the nodes' capability to store the content and forward it to other, intermittently encountered nodes?
2. What is the availability of the content, i.e. how many nodes currently carry and share the content item, and how many others have downloaded it?

The modeling approach we propose herein provides answers to the posed questions.

2.1. Measuring node interaction

Before delving into modeling, let us define relevant contact metrics. For characterizing closed systems, inter-contact time between specific node pairs has become a standard metric for representing node contact patterns. In open systems however, inter-contact times become immaterial as most node pairs have none or few contacts. We instead define contact rates as follows.

Definition 1. *Contact rate of a single node in an open system is defined as the total number of contacts that the node established during its stay normalized by its sojourn time, i.e. the number of contacts per unit time.*

The mean contact rate of the system, c, is obtained from the individual contact rates measured over a longer time interval. In particular, we denote by c_N the mean contact rate measured in a system with the average population of N nodes.

Definition 2. *Consider a snapshot of the system when the population equals N. The total contact rate Λ_N is given by $\Lambda_N = \frac{N}{2} c_N$.*

Definition 3. *Pairwise contact rate η is the rate at which two arbitrary nodes come into contact with each other.*

The total contact rate alternatively can be represented via the rate η_N, as $\Lambda_N = \frac{N(N-1)}{2}\eta_N \approx \frac{N^2}{2}\eta_N$. Note that we use index N to indicate that rates Λ_N and η_N depend on the current population size.

2.2. Stochastic model

To study phenomena such as *content survival* and to find the probability of this event, or to estimate the size of population that will carry the content, we can observe a stochastic system comprising three populations: nodes that currently carry and share the content, nodes that still have not obtained the content and nodes that stopped spreading the content. The evolution of this stochastic system can be analysed by considering stochastic processes of node arrivals and departures and the contact process between nodes. However, multivariate stochastic systems often do not easily lend themselves to analysis, due to multiple variables (in our case three) and many interacting factors that drive transitions between states. A common approach for studying complex system behaviour, which will also be utilized in our study, is by modeling by *stochastic differential equations* [9]. This approach has been used in mathematical epidemiology, where models are often referred to as *compartmental models*. Different population categories are called *compartments* and nodes within the same compartment are considered indistinguishable from one another with respect to their mobility and connectivity characteristics. By epidemic terminology, we denote as *susceptible* nodes who have not obtained content, *infected* are nodes who are still sharing the content, and those who have the content but are no longer participating in spreading are denoted as *recovered* nodes. Thus, in our

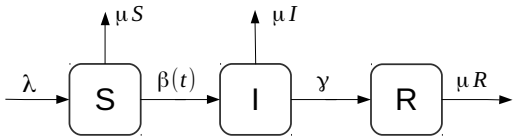


Figure 1: Division of the total population into three compartments: susceptible, infected and recovered.

model we will have compartments of: susceptible (S), infected (I) and recovered (R) nodes, Fig. 1. We will also refer to a content transfer event as an *infection*.

As a first step, we will introduce assumptions that allow us to consider the content spreading as a Markovian process. Let us assume that nodes arrive to the area according to a Poisson process with rate λ and that their sojourn time inside is exponentially distributed with mean $T = 1/\mu$. Upon arrival, all users are initially susceptible. The system state at time t is $\vec{X}(t) = [S(t), I(t), R(t)]^T$. To describe the system dynamics, we first need to consider all possible changes from state $\vec{X}(t)$ to state $\vec{X}(t + \Delta t)$ in a small time interval Δt. These changes, denoted by $\Delta\vec{X}(t)$ are given in Table 1. During Δt, a new (susceptible) node can join the system, which results in arrival to the S compartment with probability $\lambda\Delta t$. The total departure rates from the compartments are $\mu S(t)$, $\mu I(t)$, and $\mu R(t)$. We assume that the content item is small enough such that it can be transferred in zero time. Probability $\beta(t, \vec{X}(t))\Delta t$ defines an encounter event when a susceptible node becomes infected.

The rate of infection, $\beta(t, \vec{X}(t))$ depends on the number of currently active content spreading nodes $I(t)$, the number of susceptible nodes $S(t)$, and the pairwise contact rate η (defined in Section 2.1). Denote by $N(t)$ the total number of users in the area, $N(t) = S(t) + I(t) + R(t)$. In a dynamic system where population is subjected to frequent changes, the contact rate c depends on the population size: as it has been shown in [10], this dependence is linear for a certain range of population sizes. Therefore, we will consider time-dependent rates $c(t) = c(N(t))$ and $\eta(t) = \eta(N(t))$. We assume homogeneous mixing in the

Table 1: Population changes in a small time interval Δt.

Possible change	Probability
$(\Delta\vec{X}(t))_1 = [1, 0, 0]^T$	$p_1(t) = \lambda\Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_2 = [-1, 0, 0]^T$	$p_2(t) = \mu S(t)\Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_3 = [-1, 1, 0]^T$	$p_3(t) = \beta(t, \vec{X}(t))\Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_4 = [0, -1, 0]^T$	$p_4(t) = \mu I(t)\Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_5 = [0, -1, 1]^T$	$p_5(t) = \gamma I(t)\Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_6 = [0, 0, -1]^T$	$p_6(t) = \mu R(t)\Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_7 = [0, 0, 0]^T$	$p_7(t) = 1 - \sum_{m=1}^{6} p_m(t) + o(\Delta t)$
$(\Delta\vec{X}(t))_8 \neq (\Delta\vec{X}(t))_{i=1,\ldots,7}$	$p_8(t) = o(\Delta t)$

population and for the system in state $\vec{X}(t)$ we have $\beta(t, \vec{X}(t)) = S(t)I(t)\eta(t)$. After substituting $\eta(t) \approx \frac{c(t)}{N(t)}$, the rate at which susceptible nodes become infected reads

$$\beta(t, \vec{X}(t)) = S(t)I(t)\frac{c(t)}{N(t)}. \tag{1}$$

When a node receives the content item, it sets a timer for how long it will continue sharing content. Assuming an exponential timers with mean infective time $T_E = 1/\gamma$, the total recovery rate of infected nodes at time t is $\gamma I(t)$. Note that the first infected node (publisher) is assumed to use the same timer. Now that we have all transition probabilities, we can construct the SDE model. We will use one of the methods described in [11]. The following SDE system describes the dynamics of the content spreading:

$$\begin{cases} dS(t) = [\lambda - \beta(t, \vec{X}(t)) - \mu S(t)]dt + \vec{g}_1 d\vec{W}(t) \\ dI(t) = [\beta(t, \vec{X}(t)) - (\mu + \gamma)I(t)]dt + \vec{g}_2 d\vec{W}(t) \\ dR(t) = [\gamma I(t) - \mu R(t)]dt + \vec{g}_3 d\vec{W}(t) \\ \vec{X}(0) = [S(0), E(0), R(0)]^T \end{cases} \tag{2}$$

where $\vec{W}(t) = [W_i(t)]^T_{i=1,..,6}$ is a vector of six independent Wiener processes and $\vec{g}_{i=1,2,3}$ is the i-th row of the matrix

$$G(t, \vec{X}(t)) = \begin{bmatrix} \sqrt{\lambda} & \sqrt{\mu S(t)} & -\sqrt{\beta(t, \vec{X}(t))} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\beta(t, \vec{X}(t))} & -\sqrt{\gamma I(t)} & \sqrt{\mu I(t)} & 0 \\ 0 & 0 & 0 & \sqrt{\gamma I(t)} & 0 & \sqrt{\mu R(t)} \end{bmatrix}. \tag{3}$$

The SDE system (2) can be rewritten in the form:

$$\begin{cases} d\vec{X}(t) = \vec{f}(t, \vec{X}(t))dt + G(t, \vec{X}(t))d\vec{W}(t) \\ \vec{X}(0) = [S(0), E(0), R(0)]^T \end{cases} \tag{4}$$

where the vector $\vec{f}(t, \vec{X}(t))$ is given by

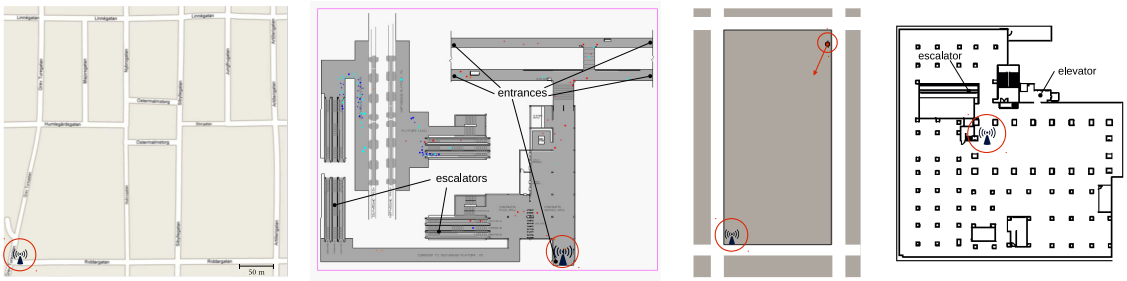
$$\vec{f}(t, \vec{X}(t)) = \begin{bmatrix} \lambda - \mu S(t) - \beta(t, \vec{X}(t)) \\ \beta(t, \vec{X}(t)) - (\gamma + \mu)I(t) \\ \gamma I(t) - \mu R(t) \end{bmatrix} \tag{5}$$

At the time when the content was first published there was a single infected node in the area and $S(0) = S_0$ susceptible nodes, thus the initial system state is $\vec{X}(0) = [S_0, 1, 0]^T$.

The solution to (4) gives the probability distribution $p(t, \vec{x})$ which satisfies, [11]:

$$\frac{dp(t, \vec{x})}{dt} = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{\partial^2}{\partial x_i \partial x_j}\left[p(t, \vec{x})\sum_{k=1}^{6} g_{i,k}(t, \vec{x})g_{j,k}(t, \vec{x})\right] - \sum_{i=1}^{3}\frac{\partial[p(t, \vec{x})f_i(t, \vec{x})]}{\partial x_i} \tag{6}$$

Above, f_i is the i-th entry of $\vec{f}$ and $g_{i,j}$ is the i, j entry of the matrix G. Eq. (6) is known as the forward Fokker-Planck equation describing the time-evolution of $p(t, \vec{x})$.



(a) (b) (c) (d)

Figure 2: Topology of the simulation areas for: Östermalm (a), Subway (b), Open RWP (c), Conference scenario (d).

3. Model validation and analysis

3.1. Mobility scenarios

Having derived the model which is based on rather favorable assumptions (from the modeling perspective), we want to investigate how well it fits more realistic mobility. To cover diverse cases, we consider four open system scenarios: two scenarios with simulator-generated traces, a synthetic mobility model and a real-life trace. Herein we give the most important details about the scenarios.

1. *Legion mobility traces*: For emulating realistic mobility, we use traces generated with *Legion Studio*[1], a multi-agent mobility simulator commonly used for pedestrian traffic planning and public spaces design and dimensioning.

We recreate two scenarios[2]. The first scenario represents an outdoor urban space, modeling a part of the downtown Stockholm, which we will further refer to as the Östermalm scenario. The topology is represented by a grid of streets, Fig. 2(a). There are fourteen streets (and eleven intersections) from where pedestrians enter the area and depart; while inside the area, they are constantly moving. The second scenario captures the movement of passengers in a subway station. The simulation area comprises two levels: the upper, entry level and the lower level with train platforms, which are connected by escalators, Fig. 2(b). Nodes can arrive either by walking in from one of the five entrance points on the upper level, or when a train arrives at the platform. The mobility model incorporates realistic behavior: nodes pausing in the coffee-shop and the store at the entry level, queuing on the escalators, and congregating at the platforms. A detailed description of these two scenarios can be found in [10].

2. *Open random waypoint (RWP)* [12] models nodes traversing a small confined area such as a city square. The main difference between this form of random waypoint and the traditional one is the varying population. Nodes enter the area from one of the entrance points according to a Poisson process and immediately start moving towards a randomly chosen point inside. Upon reaching a waypoint, the node decides either to leave the area, with probability P_{exit},

[1]Legion Studio software `http://www.legion.com/legion-studio`.
[2]Traces are available at `http://crawdad.org/kth/walkers`.

Table 2: Scenario parameters: Distributed content sharing.

	Östermalm	**Subway**	**Open RWP**	**Hope**
λ [s^{-1}]	0.43	1.04	0.20	0.062
T [s]	354	195	261	965
T_E [s]	150	60	150	200
c_N [s^{-1}]	0.102	1.04	0.070	0.102

by choosing (uniformly in random) one of the exits as its next waypoint, or chooses a new waypoint inside. The area in our setup is rectangular, with size 100 m by 200 m and there are four entrances (exits), one at each corner, Fig. 2(c). The exit probability is 0.75. Nodes travel at speed of 1 m/s, without pausing between two consecutive waypoints.

3. *Conference scenario:* To create an open, trace-based mobility model, we use mobility traces collected at the HOPE conference in 2008 [13]. The traces contain positioning data of attendees roaming on two floors of the conference hotel. We focus only on the floor where attendees exhibited higher mobility (roaming between a registration desk, exhibition area, commercial stands and similar). By extracting an hour-long sample, we obtain mobility details of 233 attendees. Fig. 2(d) depicts the floor plan and the locations of the elevator and escalators, where attendees could access or exit the floor.

3.2. Evaluation

In this section we validate our model via simulations, and we analyze the results and the model's applicability and limitations. To simulate the mobility scenarios, we use the ONE simulator [14]. In the simulations, we limit the transmission range for Östermalm and Open RWP to 10 m and for Subway and Hope to 5 m, assuming that communication in outdoor scenarios has longer range as compared to indoor scenarios. Since the time granularity of the Hope dataset is 30 s, the actual node positions between two consecutive snapshot locations were interpolated as if the nodes were moving at a constant speed. In the Hope trace, some users are already in the system at time $t = 0$, while in the other three scenarios, initially, the system is empty. For computing the scenario parameters (listed in Table 2), we discard measurements from the warm-up period, whose length we determine by applying Welch's method as described in [15]. We also use the calculated length of the warm-up period as a starting time for simulations, since we assume the system was in the steady state when the content was published.

First, we look into the average arrival rates for susceptible users, λ. The Östermalm trace was generated such that both ends of each intersecting street[3] feed the area with Poisson arrivals with rate of 0.04 s^{-1}. Likewise, the Open RWP generated arrivals with rate 0.05 s^{-1} at each of the corners. From Little's law, which for the average number of nodes reads $\bar{N} = \lambda T$, we compute the arrival rates: 1.04 and 0.062 s^{-1} for the Subway and Hope scenario, respectively. Table 2 details other parameters: the average sojourn times T and the contact rate c_N.

[3]Note that there are eleven entrance points and fourteen streets.

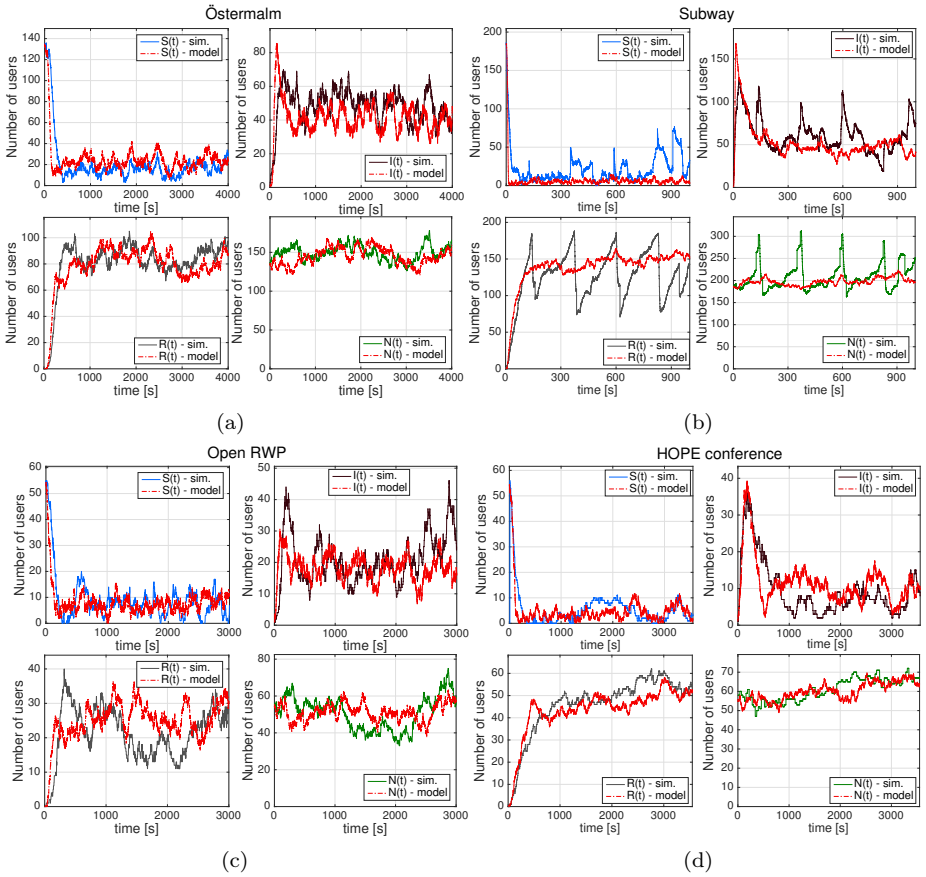


Figure 3: Simulated content sharing process and one sample path of the stochastic model: Östermalm (a), Subway (b), Open RWP (c), Conference scenario (d).

We simulate the stochastic process $\vec{X}(t)$ by plugging the computed values into Eq. (4) and plot the results in Figs. 3 and 4. The system evolution over time is plotted in Fig. 3 and we observe that the modeled processes behave similarly to those simulated from the traces in the Östermalm, Open RWP and Hope scenarios. Our model accurately predicts the initial content spreading phase with respect both to the size of the infected population and the time it takes to reach that level of infection. The model also provides a good estimation for the size of stochastic fluctuations; this is an improvement from a deterministic approximation of the process, which would give only steady state population distributions. In the Subway case, the model captures the decreasing trends in the number of susceptible nodes, as well as the increase of the recovered node population, but is unable to account for the burstiness of node arrivals and departures when a train arrives.

Next, we compare the simulations with the analytic results with respect to the estimated number of nodes of each type. Fig. 4 shows the median values

(central marks in the boxes), as well as the 25th and 75th percentiles (edges of the boxes), and the distribution outliers (cross marks) for the numbers of susceptible, infected and recovered nodes, and the total population size. The black boxes represent the simulation results and the blue ones represent the model. The model yields accurate estimations in all scenarios, slightly overestimating the number of susceptible nodes for Östermalm and Hope trace (which is still useful as a prediction of the performance lower bound).

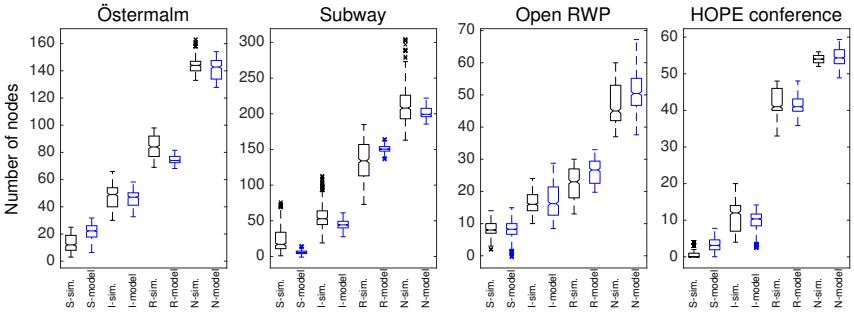


Figure 4: Susceptible, infected, recovered and the total number of users: simulation results (black) compared with the analytic results (blue).

3.3. Model limitations

Since the model assumes constant arrival rate λ, batch arrivals, such as those manifested in the Subway scenario, cannot be captured. Large batch arrivals (and departures) of size 30 to 50 users are frequent: they can be identified as peaks in Fig. 3(b). Although these short-lived events have an impact on variations of node population distributions, we can still obtain a reasonable estimation of the average numbers of each node type (measured over several minutes). In cases where variable arrival rates had to be taken into account, such requirement would modify Eq. (2) to allow for time-dependent $\lambda(t)$ (in the Subway case evidently periodic). This is however out of the scope of this study.

To summarize our first findings, we have shown that our model efficiently counts users of different types in a varying population, and can be applied in various mobility scenarios. The model relies on the Markovian property for user sojourn times and the assumption of the Poisson arrival process—such assumptions are the norm for analytical work, since modeling based on looser assumptions easily gets more complicated or even intractable. Nevertheless, our model still yields a reasonably good match with the simulations. To investigate how justified the previous assumptions are, in the next section we take a closer look at the properties of the examined traces.

3.4. Revisiting modeling assumptions

3.4.1. Contact rate distribution

We use the Kolmogorov-Smirnov (K-S) statistical test to compare the empirical contact rate distributions with standard probability distributions: log-logistic, logistic, log-normal, Gamma and Weibull distribution. The parametriza-

Table 3: K-S statistics for distributions fitted with empirical contact rate distributions

Scenario	Distribution	Value
Östermalm	Log-logistic	0.018
Subway	Log-logistic	0.019
	Log-normal	0.015
Open RWP	Log-logistic	0.038
	Logistic	0.024
	Weibull	0.038
Hope	Log-logistic	0.051
	Logistic	0.090
	Gamma	0.086
	Weibull	0.079

tion of the reference distributions is based on the maximum-likelihood estimation and the results are given in Table 3 (lower value of K-S statistic indicates better fit). Fig. 5 shows only those distributions for which the null-hypothesis was not rejected by the K-S test. Log-logistic distribution proves to be a good fit for all empirical distributions. This distribution features a narrower peak and a heavier tail than the other tested distributions, which signalizes that nodes exhibit similar connectivity properties—a large fraction can be regarded as ho-

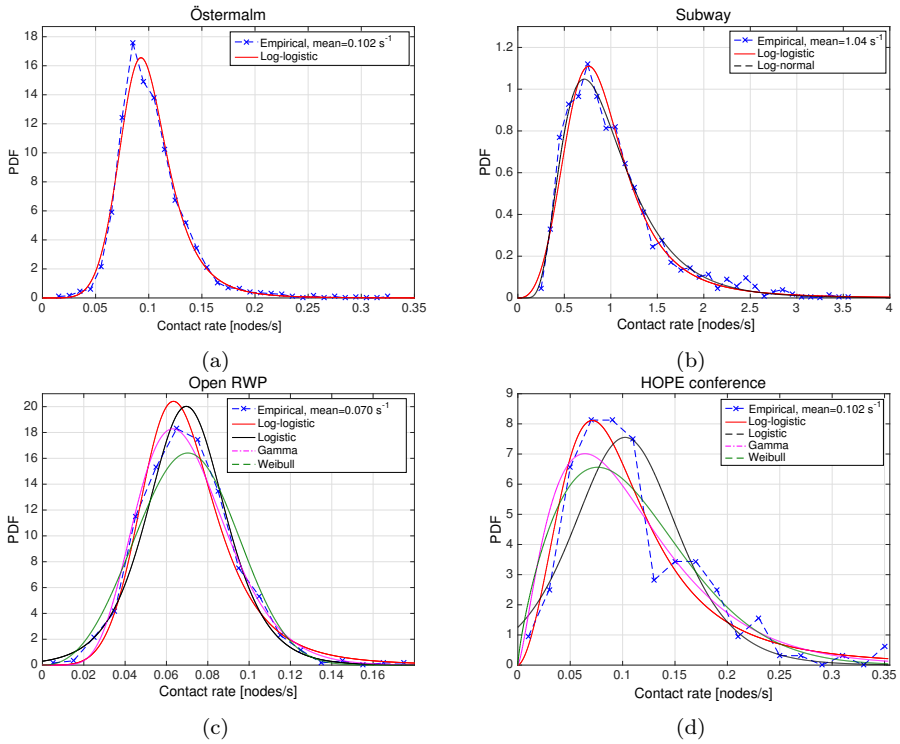


Figure 5: Contact rate distributions: Östermalm (a), Subway (b), Open RWP (c), Conference scenario (d).

mogeneous to some extent—but there are also "more social" nodes with higher contact rates.

3.4.2. Homogeneous mixing

The assumption on homogeneous mixing implies that all susceptible nodes are equally likely to be in potentially infective contact. Such assumption is obviously too lax in the beginning of the epidemic process: nodes that establish more contacts are more likely to contract the infection. However, after the content has been present in the population for some time, nodes' mobility will contribute to more spatially uniform spread of content. This can be recognized as a drop in the number of susceptible nodes (Fig. 3), which is sharper for the model-generated path than the for simulation results.

3.4.3. Sojourn time distribution

Fig. 6 depicts the sojourn time distributions for the four scenarios. Clearly, only in the Hope trace node sojourn times appear to be exponentially distributed. K-S test with the the statistics of $k = 0.055$ verified the assumption of exponentially distributed sojourn times. For other scenarios, the figure also shows hypothesized distributions, which however were all rejected by the K-S

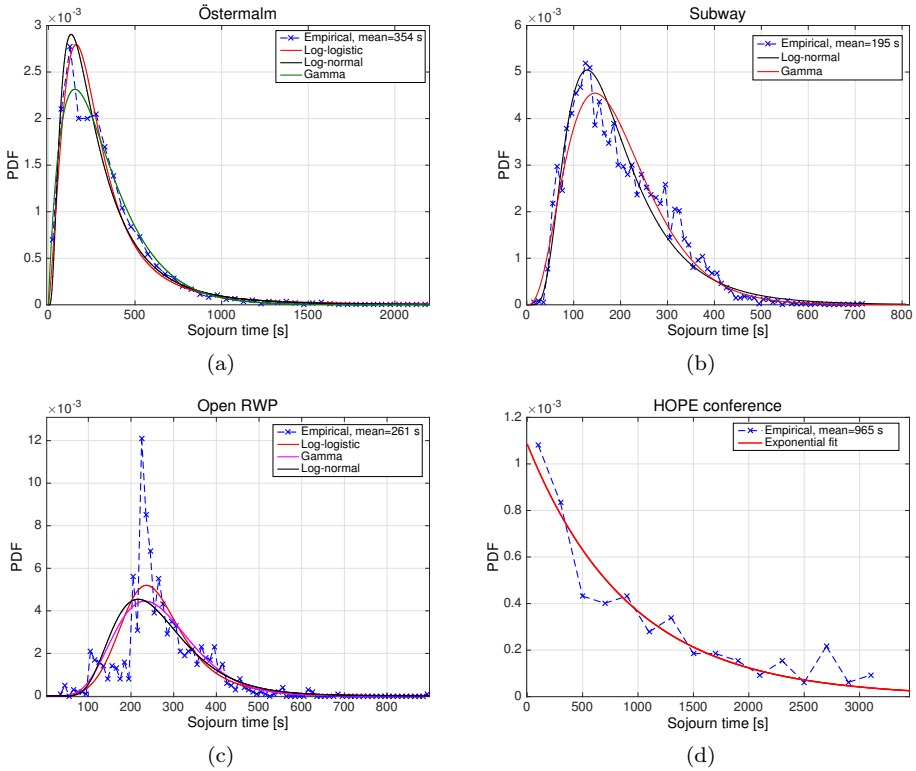


Figure 6: Distribution of node sojourn time: Östermalm (a), Subway (b), Open RWP (c), Conference scenario (d).

test. Empirical distributions implicate that nodes stay longer than assumed, which results in a higher number of recovered nodes.

3.4.4. Node heterogeneity

In previous sections, we observed that nodes exhibit heterogeneous connectivity, however, in the investigated scenarios, the average metrics seem to be sufficient to describe the system behaviour. This may not be the general case. In particular, when there is a significant variability in contact rates, contacts between infected and susceptible individuals will initially be more frequent than predicted by homogeneous-mixing models.

3.4.5. Discussion

Since our scenarios are characterized by moderate node heterogeneity, the model has not exposed itself as sensitive to exact distributions and their deviations from Markovian assumptions. We have noticed however, that in order to achieve good fit with the simulation results, it is of utmost importance to accurately estimate the parameters, particularly the contact rate. Given the large number of nodes in all scenarios, inaccuracy in contact rate estimation will result in the miscalculation of the infection rate, which is essential for predicting the course of the content spreading, as well as the final sizes of different sub-populations.

3.5. Probability of content survival

For the content to survive in the area, it is crucial that the publisher first meets enough nodes and replicate the content, and that, subsequently, other infected nodes continue spreading it. Due to the probabilistic nature of node contacts, it may happen that all nodes with content either leave the area or stop spreading the content; in that case no further spreading is possible and the content is lost. Probability of such events depends on the frequency of contacts, as well as the number of nodes in the area (occupancy) and the infectious period. However, contact rate and occupancy cannot be engineered. The only remaining parameter is the length of the infectious period T_E, that is, the time that nodes are willing to allocate their resources in support of content spreading. In the most generous case T_E would equal the node's sojourn time. The objective of a node, expectedly, is to minimize its own cost, while still supporting the system. Thus, for the contents to survive, or to achieve a certain level of availability measured as the number of nodes that have obtained the content, T_E should be chosen appropriately.

In a deterministic model, content spreading definitely occurs as soon as there is at least one node with the content. In a stochastic model, the content may disappear before it reaches enough nodes to support new infections. We are interested in finding the *probability of content survival*. Although there are available methods for approximating the spreading process during its early stage (e.g. as a birth-death process), for our model such approximations become non-trivial, due to the state-dependent infection rate and existence of recovered

nodes. We resort to simulating the process $X(t)$ to find the empirical probability of content survival, plotted in Fig. 7 (we show only results for the Östermalm case). At a certain threshold for the timer T_E^{th} (around 100 s), the probability of survival reaches 0.8 and increases only marginally beyond that threshold. Thus, above the threshold T_E^{th}, content is equally likely to survive, whereas the availability of content still depends on the value T_E.

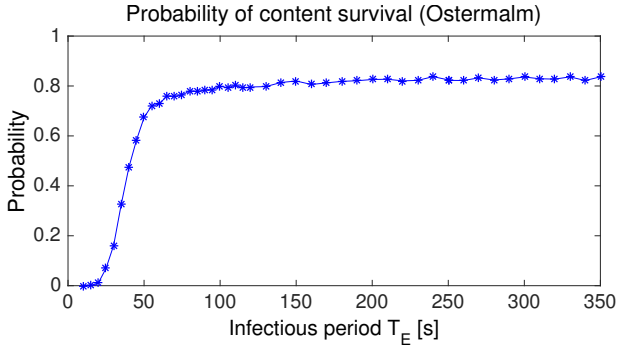


Figure 7: Probability of content survival as a function of the infectious period.

4. Opportunistic content sharing with infrastructural support

Earlier, we have assumed that the system parameters (user arrival rates, sojourn times and contact rates) were known. In this section we present an example application, a tool which can be used to gather mobility and connectivity data about users in the observed area. We do not address the technical details of the application, but merely propose the concept of how the data collection can be realized.

The communication scheme we consider is infrastructure-supported opportunistic content spreading. The setup is similar to the distributed case in Section 2: nodes arrive to the area, spend some time inside and eventually leave; additionally, a stationary node—access point—is deployed in the area. The publisher here is the access point, which is able to reach only a fraction of the mobile population that can be found in the access point's vicinity. Nodes that obtain contents from the access point, due to their mobility and by means of opportunistic communication, extend the content availability in the rest of the area. Further, mobile nodes are able to determine their approximate position (e.g. from GPS coordinates or by WiFi triangulation methods).

The access point publishes content of relevance for larger audience (e.g. weather updates, news, or announcements). We can assume that all nodes traversing the area are interested in obtaining content, which is done through an application that also collects logs about content dissemination. The application imposes little intrusion on the user privacy, collecting only the timestamps when the content was downloaded and when the user departed from the area, as well as the number of encountered users. The content itself can be regarded as accessible for download only within the locale, which could be one of the ways to incentivize users to collect and contribute their measurements.

Table 4: Population changes in a small time interval Δt.

Possible change	Probability
$(\Delta\vec{X}(t))_1 = [1,0,0]^T$	$p_1(t) = \lambda_p \Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_2 = [-1,0,0]^T$	$p_2(t) = \mu_p P(t) \Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_3 = [0,1,-1]^T$	$p_3(t) = \beta(t, \vec{X}(t)) \Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_4 = [0,-1,0]^T$	$p_4(t) = \mu_u S(t) \Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_5 = [0,0,1]^T$	$p_5(t) = \lambda_u \Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_6 = [0,0,-1]^T$	$p_6(t) = \mu_u U(t) \Delta t + o(\Delta t)$
$(\Delta\vec{X}(t))_7 = [0,0,0]^T$	$p_7(t) = 1 - \sum_{m=1}^{6} p_m(t) + o(\Delta t)$
$(\Delta\vec{X}(t))_8 \neq (\Delta\vec{X}(t))_{i=1,\ldots,7}$	$p_8(t) = o(\Delta t)$

The data gathering application operates as follows. Whenever a new user arriving to the area passes by an access point, which is located at some of the entrance points, the user will receive the content. This user will be denoted a *primary infected* user. The application on the access point logs user details with the time of arrival and the location of the node's entrance point. While infected users roam inside the area, they will meet uninfected users and forward the content. Upon such event, the uninfected become *secondary (infected)* users[4]. The application on the newly infected node then sends to a server (e.g. via cellular network) timestamp of the new infection event. There will be a number of users who do not meet any infected users during their stay, and eventually leave the area without finding the content. We label these as *unreached* users (also *uninfected*). Infected users, primary and secondary, will continue spreading the content, and count the users they encounter—both infected and previously uninfected. Finally, when an infected user leaves the area, the user application again uploads information about the user: the exit location, departure time and the measured number of encounters; these new data is matched with the time of infection for each infected user. Optionally, the application can periodically send collected measurements while the user is still in the area. As we will describe later, the encounter information will be used to compute contact rate between users.

With this method, we can also answer the following:

1. What is the number of users currently residing inside the area, and how does this number vary over time?
2. How many users cannot be reached by this content sharing scheme?

Unlike in the purely opportunistic case, in this scenario there is no risk of irreversibly losing the content. As long as there is an inflow of infected users to the area, if the content vanishes from the area, i.e. when all infected users have left the area, the content will be re-injected with an arrival of a user infected by an access point.

[4]We will interchangeably use the short and the full denotation.

4.1. Model

We want to estimate the number of users, which can be one of the three aforementioned types, their sojourn times and contact rate of users in the areas. Denote by $P(t)$ the number of primary infected users, by $S(t)$ the number of secondary infected[5], and by $U(t)$ the number of unreached users at time t. Let us assume that primary users arrive to the area according to a Poisson process with rate λ_p and that their sojourn time inside is exponentially distributed with mean $T_p = 1/\mu_p$. Assume also that the total arrival rate of all other (non-primary) infected users is λ_u. The sojourn time of both the S and U type comes from the same distribution with mean value $T_u = 1/\mu_u$[6]. Then, arrivals to the area constitute a Poisson process with the total rate $\lambda_p + \lambda_u$ and all node sojourn times are exponentially distributed.

The system state at time t is $\vec{X}(t) = [P(t), S(t), U(t)]^T$; it is not fully observed since we do not know the number of unreached users $U(t)$. We can however infer their approximate values by modeling the system dynamics, governed by stochastic processes which include arrivals, departures and transitions from uninfected to secondary infected users. Transitions between the compartments, as well as the external arrivals and departures in the model are illustrated in Fig. 8.

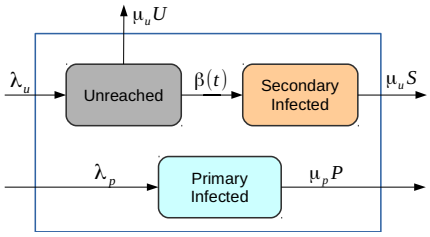


Figure 8: Compartmental model with transition rates.

Similarly as in Section 2.2, we find all possible state transitions in Δt (Table 4). An arrival can occur either in the P or U compartment with probabilities of these events $\lambda_p \Delta t$ and $\lambda_u \Delta t$, respectively. The total departure rates from the three compartments are $\mu_p P(t)$, $\mu_u S(t)$, and $\mu_u U(t)$. An unreached node becomes secondary infected with rate $\beta(t, \vec{X}(t))$ (transition $U \to S$). This is, again, the rate of infection, which herein is a function of the total number of infected users, $I(t) = P(t) + S(t)$, the number of unreached users $U(t)$, and the contact rate $c(t)$, $\beta(t, \vec{X}(t)) = I(t)U(t)\frac{c(t)}{N(t)}$.

Following a similar approach as in Section 2.2, we get the SDE system:

[5]Remark that S here stands for (secondary) infected users, not for susceptible users as in Section 2.

[6]For the purpose of modeling, we allow different sojourn times T_p and T_u, but it is justifiable to assume $T_p = T_u$.

$$\begin{cases} dP(t) = [\lambda_p - \mu_p P(t)]dt + \vec{g}_1 d\vec{W}(t) \\ dS(t) = [\beta(t, \vec{X}(t)) - \mu_u S(t)]dt + \vec{g}_2 d\vec{W}(t) \\ dU(t) = [\lambda_u - \beta(t, \vec{X}(t)) - \mu_u U(t)]dt + \vec{g}_3 d\vec{W}(t) \\ \vec{X}(0) = [P(0), S(0), U(0)]^T \end{cases} \tag{7}$$

keeping the same denotation for $\vec{W}(t)$ and $\vec{g}$ as in Eq. (2), whereas the matrix $G(t, \vec{X}(t))$ is given as

$$G(t, \vec{X}(t)) = \begin{bmatrix} \sqrt{\lambda_p} & \sqrt{\mu_p P(t)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\beta(t, \vec{X}(t))} & \sqrt{\mu_u S(t)} & 0 & 0 \\ 0 & 0 & -\sqrt{\beta(t, \vec{X}(t))} & 0 & \sqrt{\lambda_u} & \sqrt{\mu_u U(t)} \end{bmatrix}. \tag{8}$$

The initial state $\vec{X}(0)$ is (partially) known. Expression (7) can be shorten as:

$$\begin{cases} d\vec{X}(t) = \vec{f}(t, \vec{X}(t))dt + G(t, \vec{X}(t))d\vec{W}(t) \\ \vec{X}(0) = [P(0), S(0), U(0)]^T \end{cases} \tag{9}$$

where the vector $\vec{f}(t, \vec{X}(t))$ is now given by

$$\vec{f}(t, \vec{X}(t)) = \begin{bmatrix} \lambda_p - \mu_p P(t) \\ \beta(t, \vec{X}(t)) - \mu_u S(t) \\ \lambda_u - \beta(t, \vec{X}(t)) - \mu_u U(t) \end{bmatrix}. \tag{10}$$

Note that although Eqs. (2) and (7) appear to be similar, the main insights into the system behaviour can be obtained from the Eq. (6).

5. Evaluation

Now we validate the second model and in addition to the previously used traces, we use one more real-life trace.

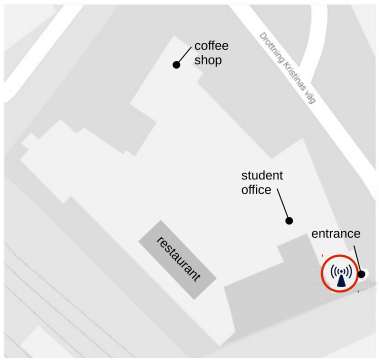


Figure 9: Simulation area for the Campus scenario.

Campus scenario: This dataset comprises mobility traces collected from a campus-wide WLAN. We concentrate on mobility inside a student union building (Fig. 9) and extract 1.5 hours of associations. The wireless network inside the building comprises of 28 access points located on three floors. The sampled trace captures movement of total 442 wireless network users over the observed period, with the average occupancy of 82 users. Users can leave the area and return later: we consider such users as new arrivals, rather than returning users. Since we do not have exact positions, nor the contact traces of users in this set, we assume that two users are able to establish contact when associated with the same access point.

5.1. Model validation

In each scenario, there is a single access point located as depicted in Fig. 2 (a)–(d) and Fig. 9. We keep the same transmission ranges as in the previous simulations (10 m for indoor and 5 m for outdoor scenarios). Access points have the same transmission range as mobile nodes. The area of the Campus trace is a multi-floor building. When simulating this scenario, we choose an access point that is located near one of the entrances.

We first estimate the average arrival rates for primary users, λ_p. In the Östermalm and Open RWP trace these rates represent arrivals at one of the entrance points, thus their values are 0.04 s^{-1} and 0.05 s^{-1} (see Section 3.2). Applying Little's law to the average number of primary users $\bar{P} = \lambda_p T_p$, we compute the arrival rates: 0.195, 0.012 and 0.010 s^{-1} for the Subway, Hope and Campus scenario. Similarly, from the simulation results we find the arrival rates for non-primary users λ_u: 0.424, 0.956, 0.150, 0.049 and 0.072 s^{-1}. When computing the contact rate c_N, in order to achieve higher confidence, we utilize the encounter information both from primary and secondary infected users. The contact rate measured by a secondary user is then the number of encounters normalized with its sojourn time in the S state. Note that in the model validation step all parameters were computed offline. Parameters for the five scenarios are summarized in Table 5.

It can be interesting to study the stochastic process in Eq. (9) in the transient state. In three scenarios, the system starts from an empty state; thus, before reaching the steady state, the average contact rate will be lower than c_N. Since we can measure only the contact rate c_N, we will introduce a heuristic approximation $c(t) = c_N \frac{N(t)^2}{2N^2}$ for the contact rate in the transient state, while

Table 5: Scenario parameters: Infrastructure-supported content sharing.

	Östermalm	**Subway**	**Open RWP**	**Hope**	**Campus**
λ_p [s^{-1}]	0.040	0.195	0.050	0.012	0.010
λ_u [s^{-1}]	0.424	0.956	0.150	0.049	0.072
T_p [s]	360	188	266	1016	489
T_u [s]	307	191	259	954	980
c_N [s^{-1}]	0.102	1.04	0.070	0.102	0.039

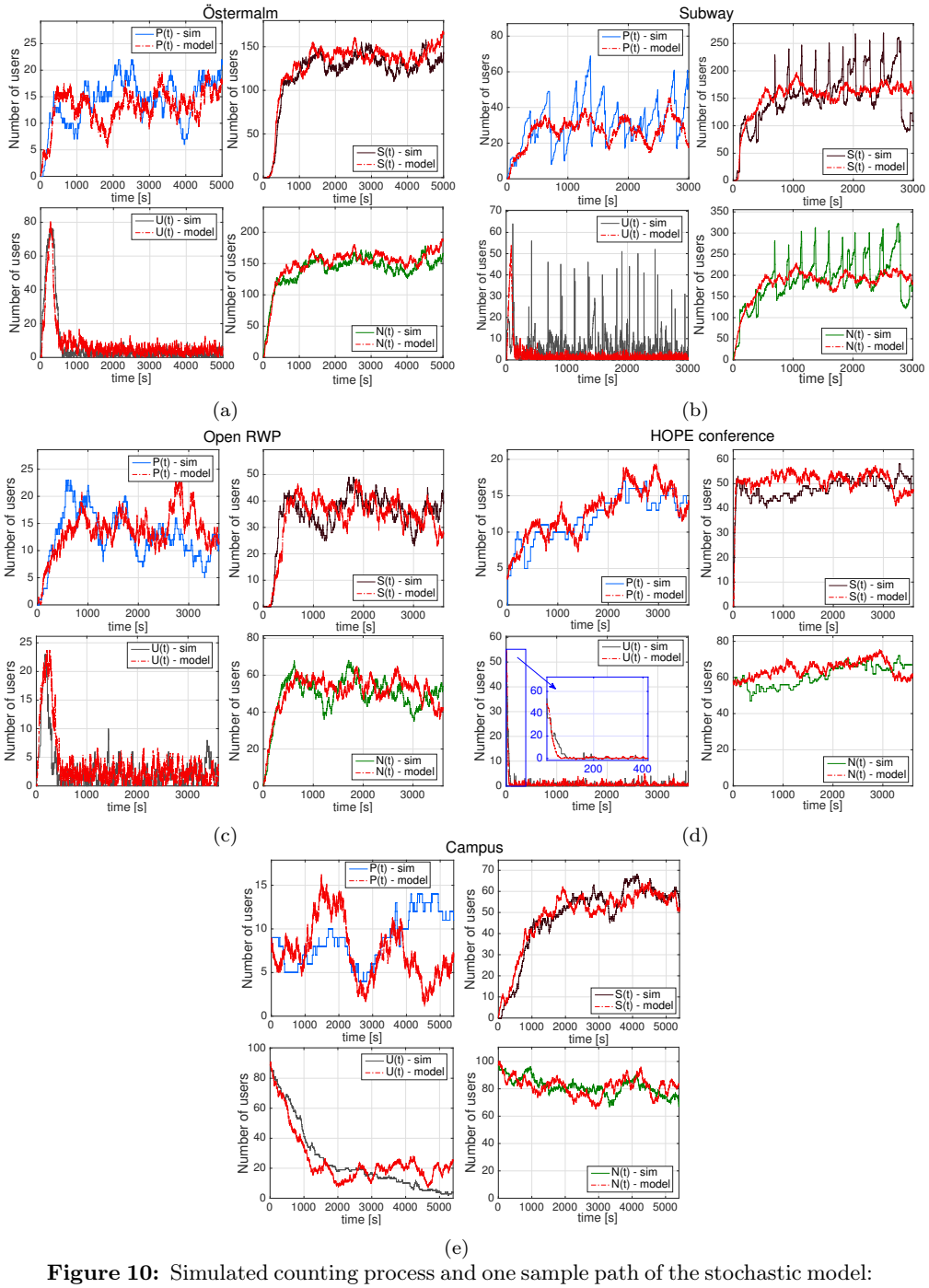


Figure 10: Simulated counting process and one sample path of the stochastic model: Östermalm (a), Subway (b), Open RWP (c), Conference scenario (d), Campus (e).

in the steady state, the contact rate is calculated as $c(t) = c_N \frac{N(t)}{N}$. Note that, in a real deployment, the population size estimation would not be possible without any prior knowledge of system parameters, e.g. the expected contact rate in the observed area. The estimation requires an initial data acquisition period, whose length depends on the arrival rates and sojourn times. Once we have obtained the contact rate for a specific setup, this estimation will be readily available for future uses: for example, if we wanted to model the occupancy of the space during different times of the day.

Now we compare the simulation results for the system evolution with the analytic predictions of the stochastic process $\vec{X}(t)$. The results are depicted in Fig. 10. For the Östermalm, Subway, Open RWP and Hope scenario we simulate realizations of the content spreading directly from Eq. (9). For the Campus scenario, after the initial tests we recognized that a minor modification in the system equation is required. Namely, we noticed that the analytic expression Eq. (7) for the Campus trace predicts much faster spread of infection than what was observed from the simulations. The cause of this discrepancy is the overly optimistic estimation of the infection rate $\beta(t, \vec{X}(t))$. Relying on the assumption of homogeneous mixing, the infection rate incorporates the possibility that a pair of any two users (where one is infected and the other is not) in the entire population currently roaming in the area, is equally likely to establish a contact. By examining the trace, we found this assumption insupportable: not just that users do not mix homogeneously, they appear to be spatially separated into groups and only some users roam from one group to another. As a result of such mixing, the infected nodes will be more likely to meet other infected users and contribute less to the spreading. Thus, the infection rate has to be scaled down to represent sparse mixing. We arrive at a new expression for the infection rate:

$$\beta(t, \vec{X}(t)) = I(t) \frac{U(t)}{\kappa} \frac{c(t)}{N(t)}, \tag{11}$$

where κ is introduced as a scaling factor. For the campus trace, we found empirically that with $\kappa = 8$, Eq. (11) approximates the infection rate observed in simulations. Returning to the Fig. 10, we observe a quite good match between the modeled and the simulated processes in the Östermalm, Open RWP and Hope scenarios. First, observe that the model accurately predicts the system behavior in the transient state, matching the duration of the transient period, and the population size changes for all user types. In the steady state prediction works equally well, except for the Subway trace were the model is again unable to cope with non-Poissonian arrivals but it accurately captures the average numbers of different user types.

Looking at the population distributions for the primary, secondary, unreached users and the total population size (Fig. 11) the model achieves good predictions in all scenarios, slightly overestimating the number of unreached users for Östermalm, Open RWP and Campus traces, and underestimating their number in the Subway case. These numbers are low in all cases, ranging up to several users: due to the high node density, in the simulations there are very few (in fact, close to zero) users who never meet anyone.

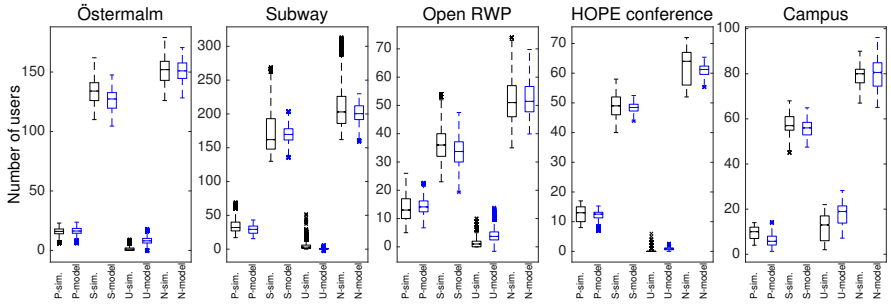


Figure 11: Primary, secondary, unreached and the total number of users: simulation results (black) compared with the analytic results (blue).

5.2. New findings

In this section, we validated another example of an epidemic-like, opportunistic content sharing scheme, which can be modeled by our SDE modeling approach. We also discovered that neglecting the heterogeneity of contacts can give misleading results. However, in certain cases such as the Campus scenario, heterogeneous mixing can be approximated by introducing a scaling parameter when calculating the infection rate. This minor modification to the model comes at a cost of estimating the best-fit scaling parameter for the scenario in hand, but it also brings the advantage of using a relatively simple and tractable model.

5.3. Application

Now that we have a better understanding of when and how the model can be applied, we revisit one of the initial objectives to estimate the number of unreached users, $U(t)$.

Since the application logs the arrival and the departure times, as well as $P(t)$, the arrival rate λ_p and the departure rate μ_p can be readily computed. Likewise, the average time a node spends as a secondary infected user can be found from the node's infection and departure timestamps. The remaining challenge is to estimate the arrival rate λ_u. To this end, we use approximate Bayesian computation (ABC) implemented in the `abc-sde` software tool [16, 17]. ABC is a "likelihood-free" methodology for Bayesian inference, particularly suitable for models with intractable or computationally demanding likelihood function, as well as partially observed models, like the one we consider. ABC takes as input a (partially observed) trajectory sample and estimates the model parameters by targeting an approximation to their posterior distributions (with respect to the observed sample). Our goal is to estimate a single parameter, λ_u. The method is not restricted to inference of a single unknown parameter, but can be used for a set of parameters; this however becomes more computationally expensive and less reliable for partially observed systems.

For the Campus scenario we use a 200 s long sample and for all other scenarios, a 150 s long sample $Y(t_n) = [P(t_n), S(t_n)]^T$, $n = 1, ..., N_s$, $N_s \in \{150, 200\}$

and run the estimation algorithm. The estimated arrival rates are: 0.329, 0.914, 0.126, 0.055 and 0.078 s^{-1}. Compare these rates with the corresponding λ_u from Table 5: while we have very accurate estimations for the Subway and the Hope scenario, in the Östermalm case underestimating λ_u results in the difference of 16 users (out of around 151 in the system) during the observed time interval. Nevertheless, an approximate estimation can be sufficient to determine the parameter space, and then iteratively improve the estimation by fine tuning. This is done by simulating the counting process $\vec{X}(t)$ with the estimated parameters, and then comparing the simulation results for the number of primary and secondary users with the measured data. In this way, it can be easily inferred whether the actual arrival rate is higher or lower than the estimated value; this inference is subsequently used to iterate simulations until the simulated process is fitted to the measurements.

6. Related Work

This paper revisits an epidemiological approach to model opportunistic message spreading, and presents new findings in opportunistic contact characterization—specifically when considering user population with churn—and we position our work with respect to these contributions.

Modeling spread of messages in mobile ad hoc networks inspired by the spread of infectious diseases between humans was first applied to epidemic routing in [18]. This approach has been further extended in [19], where the authors proposed a framework for studying a variety of epidemic routing schemes. Their common starting point is the use of ordinary differential equation models in performance evaluation. Following these studies, there has been a wealth of work considering various content spreading schemes through the deterministic approximations of Markovian chains representing content epidemics. While the deterministic models can be useful for estimating important characteristics of epidemics (e.g. conditions under which they occur, the rate at which they grow, the expected number of infections in equilibrium), such models become unreliable when the population is small, and the process exhibits stochastic fluctuations that cannot be neglected. Furthermore, ignoring the stochastic nature of an epidemic, such models are unable to provide information about other important features of the dynamics, including the size of fluctuations in the number of infected individuals and the possibility that fluctuations will result in extinction of the infection.

We resort to the field of mathematical epidemiology, and adopt the stochastic differential equation (SDE) modeling approach. There, the SDE approach has already produced a substantial amount of work, encompassing studies on both open and closed populations, different disease transmission patterns, derivation of the stability conditions and so on [20]. To the best of our knowledge, our approach is the first to propose the use of SDE models for opportunistic networks analysis.

Characterization of node interactions in open opportunistic systems is a relatively new topic. An empirical study on the impact of pedestrian mobility

on connectivity in wireless systems was presented in [10]. Implications of the population churn on the system performance have been investigated in [7, 12]; while these works consider a particular content sharing scheme, our aim is to provide models which can be adapted to various use-cases.

Heterogeneity has been recognized as one of the important features to produce realistic scenarios. We show for a range of mobility scenarios and two communication schemes that a detailed overview of nodes' contact patterns is not essential and can be represented by measuring average contact rate of nodes. We however refrain from ascertaining that this would hold in general: some routing schemes, for example, explicitly exploit node heterogeneity to find the best candidates for fast and resource-efficient delivery. Capturing contact rate variability for such purposes is advantageous. Spyropoulos et al. show in [21] how heterogeneity can be introduced by separating nodes into different communities. As we have seen, it is enough that nodes in a community share similar properties. Thus, heterogeneous population can be subdivided into a small number of distinct communities; some nodes may belong to multiple communities, acting like bridges between those.

7. Conclusion

In this paper we studied the operation of two opportunistic content sharing schemes in systems with varying population. The schemes—a completely distributed scheme and a scheme which incorporates infrastructural nodes—deploy epidemic-style content spreading. By using the stochastic differential equations modeling approach, we developed models for the operation of these schemes. We analyzed the models, validated them against both synthetic and real-life mobility traces, and we demonstrated that the models prove to be a good match for the examined scenarios. Since the scenarios have quite diverse features, we believe that the modeling approach is suitable for a broad range of applications, under certain assumptions on the user mobility behavior and connectivity properties.

References

[1] L. Pajevic, G. Karlsson, Characterizing opportunistic communication with churn for crowd-counting, in: Proc. IEEE WoWMoM AOC Workshop, 2015, pp. 1–6.

[2] A. Picu, T. Spyropoulos, Performance of distributed algorithms in DTNs: Towards an analytical framework for heterogeneous mobility, in: Proc. Global Telecommunications Conference (GLOBECOM 2011), 2011, pp. 1–6.

[3] C. Boldrini, M. Conti, A. Passarella, Performance modelling of opportunistic forwarding under heterogenous mobility, Computer Communications 48 (2014) 56–70.

[4] J. Ott, E. Hyytiä, P. Lassila, J. Kangasharju, S. Santra, Floating content for probabilistic information sharing, Pervasive and Mobile Computing 7 (6) (2011) 671–689.

[5] Ó. Helgason, E. A. Yavuz, S. T. Kouyoumdjieva, L. Pajevic, G. Karlsson, A mobile peer-to-peer system for opportunistic content-centric networking, in: Proc. ACM SIGCOMM Mobiheld Workshop, New Delhi, India, 2010, pp. 21–26.

[6] M. May, G. Karlsson, O. Helgason, V. Lenders, A system architecture for delay-tolerant content distribution, in: Proc. IEEE Conference on Wireless Rural and Emergency Communications (WreCom), 2007.

[7] E. Hyytiä, J. Virtamo, P. Lassila, J. Kangasharju, J. Ott, When does content float? characterizing availability of anchored information in opportunistic content sharing, in: Proc. IEEE INFOCOM, Shanghai, China, 2011, pp. 3137 –3145.

[8] A. A. V. Castro, G. D. M. Serugendo, D. Konstantas, Hovering information–self-organizing information that finds its own storage, in: Autonomic Communication, Springer, 2009, pp. 111–145.

[9] B. Oksendal, Stochastic differential equations: an introduction with applications, Springer Science & Business Media, 2013.

[10] Ó. Helgason, S. Kouyoumdjieva, G. Karlsson, Opportunistic communication and human mobility, IEEE Transactions on Mobile Computing 13 (7) (2014) 1597–1610.

[11] E. Allen, L. Allen, A. Arciniega, P. E. Greenwood, Construction of equivalent stochastic differential equation models, Stochastic Analysis and Applications 26 (2008) 274–297.

[12] M. Desta, E. Hyytiä, J. Ott, J. Kangasharju, Characterizing content sharing properties for mobile users in open city squares, in: Proc. IEEE/IFIP WONS, Banff, Alberta., 2013.

[13] C. Petro, CRAWDAD data set hope/amd (v. 2008-08-07), Downloaded from http://crawdad.org/hope/amd/ (Aug. 2008).

[14] A. Keränen, J. Ott, T. Kärkkäinen, The ONE Simulator for DTN Protocol Evaluation, in: Proc. SIMUTools '09, Rome, Italy, 2009.

[15] M. Prasad, I. Ricki, Evaluation of methods used to detect warm-up period in steady state simulation, in: Proc. Winter Simulation Conference, 2004, pp. 663–671.

[16] U. Picchini, Inference for SDE models via approximate Bayesian computation, Journal of Computational and Graphical Statistics 23 (4) (2014) 1080–1100.

[17] U. Picchini, Abc-sde: a Matlab toolbox for approximate Bayesian computation (ABC) in stochastic differential equation models. (2013).
URL `http://sourceforge.net/projects/abc-sde/`.

[18] T. Small, Z. J. Haas, The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way), in: Proc. ACM MobiHoc '03, ACM, New York, NY, USA, 2003, pp. 233–244.

[19] X. Zhang, G. Neglia, J. Kurose, D. Towsley, Performance modeling of epidemic routing, Elsevier Computer Networks 51 (10) (2007) 2867–2891.

[20] L. J. Allen, F. Brauer, P. Van den Driessche, J. Wu, Mathematical epidemiology, Springer, 2008.

[21] T. Spyropoulos, T. Turletti, K. Obraczka, Routing in delay-tolerant networks comprising heterogeneous node populations, IEEE Transactions on Mobile Computing 8 (8) (2009) 1132–1147.

Ljubica Pajevic is a doctoral student at the Laboratory for Communication Networks, KTH Royal Institute of Technology, Stockholm, Sweden. She received the M.Sc. degree in System engineering and radio communications in 2009 and the B.Sc degree in Telecommunications in 2007, both from the School of Electrical Engineering, University of Belgrade, Serbia. Her research interests are in the area of opportunistic networking.

Gunnar Karlsson is professor of KTH Royal Institute of Technology, since 1998 where he is director of the Laboratory for Communication Networks. He has previously worked for IBM Zurich Research Laboratory and the Swedish Institute of Computer Science (SICS). His Ph.D. is from Columbia University, New York, and he holds an M.Sc. from Chalmers University of Technology in Gothenburg, Sweden. He has been visiting professor at EPFL, Switzerland, the Helsinki University of Technology in Finland, and ETH Zurich in Switzerland. His current research relates to network architecture, quality of service, wireless LAN and opportunistic communication.