Tracking Message Spread in Mobile Delay Tolerant Networks

Manoj Panda, Arshad Ali, Tijani Chahed, Eitan Altman

▶ **To cite this version:**

Manoj Panda, Arshad Ali, Tijani Chahed, Eitan Altman. Tracking Message Spread in Mobile Delay Tolerant Networks. IEEE Transactions on Mobile Computing, Institute of Electrical and Electronics Engineers, 2015, 14 (8), pp.1737-1750. <10.1109/TMC.2014.2362746>. <hal-01205134>

HAL Id: hal-01205134
https://hal.inria.fr/hal-01205134

Submitted on 24 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tracking Message Spread in Mobile Delay Tolerant Networks

Manoj Panda, Arshad Ali, Tijani Chahed and Eitan Altman

***Abstract*—We consider a Delay Tolerant Network under two message forwarding schemes – a non-replicative *direct delivery* scheme and a replicative *epidemic routing* scheme. Our objective is to track the degree of spread of a message in the network. Such estimation can be used for on-line control of message dissemination. With a homogeneous mobility model with *pairwise* i.i.d. exponential inter-meeting times, we rigorously derive the system dynamic and measurement equations for optimal tracking by a Kalman filter. Moreover, we provide a framework for tracking a large class of processes that can be modeled as *density-dependent Markov chains*. We also apply the same filter with a heterogeneous mobility, where the *aggregate* inter-meeting times exhibit a power law with exponential tail as in real-world mobility traces, and show that the performance of the filter is comparable to that with homogeneous mobility. Through customized simulations, we demonstrate the trade-offs and provide several insightful observations on how the number of observers impacts the filter performance.**

***Index Terms*—disruption tolerant networks; intermittent connectivity; non-replicative and replicative dissemination; epidemic routing; linear estimation; Kalman filtering**

I. Introduction

Delay Tolerant Networks (DTNs) are a class of networks characterized by intermittent connectivity and relatively long delays caused by frequent link disruptions [1]. A prime example of DTNs is a *sparse* Mobile Ad hoc NETwork (MANET) in which two nodes can communicate only when they come within the radio range of each other owing to their mobility [2]. Several other examples of DTNs exist including sparse Vehicular Ad hoc NETworks (VANETs) [3], Inter-Planetary Networks (IPNs) [4], Pocket Switched Networks (PSNs) [5], Airborne Networks (ANs) [6], Mobile Social Networks (MSNs) [7], UnderWater Networks (UWNs) [8] and networks for developing regions [9].

Several methods have been proposed in the literature for data dissemination in such mobile DTNs, e.g., *direct delivery* [10], *two-hop routing* [10], *epidemic routing* [11], *probabilistic routing* [12], *single copy forwarding* [13], and *spray-and-wait routing* [14]. However, there exists a trade-off between the delivery delay and the usage of network resources; reduction of delivery delay requires more resources such as buffer space and/or transmit power. In direct delivery, for instance, the source waits until it meets with the end destination(s) and directly delivers the message(s) to the end destination(s), and thus, incurs the maximum delivery delay. Epidemic routing, on the other hand, amounts to *flooding* of the message(s) to (possibly) all nodes in the network and results in the minimum delivery delay. In terms of the usage of network resources, direct delivery (resp. epidemic routing) uses the minimum (resp. maximum) number of transmissions.

In recent years, a novel framework has evolved to address the above trade-off [15], [16], [17], [18], [19]. The idea is to first develop a deterministic *fluid-limit* model of the network evolution under an appropriate form of open-loop control, which provides the *state trajectory*, and then apply optimal control theory to derive the optimal *control trajectory*. The same framework has also been applied for optimal containment of and recovery from malware [20], [21], [22]. Fluid-limit models, sometimes referred to as *mean-field* models,[1] have also been applied to several other networking contexts [23], [24], and [25].

The need for developing a deterministic fluid-limit model arises due to the fact that *there exist a large number of nodes in the network, and hence, it is practically impossible to work with an exact stochastic model*. The simplicity of the fluid-limit models, however, comes at the price of losing the information about the randomness in the system. A better approach is to work with a diffusion model, which provides a compromise between a deterministic fluid-limit model and an exact stochastic model. A diffusion model approximates (in a certain sense) the randomness in the original stochastic process, often called the *process noise*, in terms of Brownian motions that are often easier to work with than the original stochastic process. Diffusion models have been applied in the context of P2P networks [26], [27], [28], [29], [30] and multicasting [31].

In almost all of the above applications, one can obtain some kind of feedback on the degree of spread of messages or malware or content, etc., possibly, after some delay and/or with corruption by some *observation noise*. One might be able to specify the feedback or observation process by some function of the random state trajectory and the observation noise. The availability of diffusion models for the system dynamics and

Manoj Panda is with the Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia.
E-mail: mpanda@swin.edu.au
Arshad Ali is with the Computer Science & IT Department, University of Lahore, Pakistan.
E-mail: arshad.ali@cs.uol.edu.pk
Tijani Chahed is with the Networks & Services Department, Telecom SudParis, France.
E-mail: tijani.chahed@it-sudaris.eu
Eitan Altman is with the MAESTRO Group, INRIA Sophia Antipolis, France.
E-mail: eitan.altman@inria.fr

[1]Mean-field models are fluid-limit models where the scaling is performed w.r.t. the number of agents/entities.

that for the observation process can facilitate the application of stochastic optimal control theory with feedback to control specific realizations or sample paths in the same way as the availability of a fluid model for the system dynamics facilitates the application of optimal control theory for deterministic open-loop control.

In fact, for *linear-quadratic-Gaussian* (LQG) problems, that is, if the system and observation equations are linear and the optimization objective is quadratic (w.r.t. the state and control variables), and the system and observation noise are Gaussian, then the solution to the stochastic control problem can be obtained by estimating the state by *Kalman filtering* [32] and applying it to the corresponding deterministic optimal control problem. This is called the *separation principle* [33].

Motivation: The above discussion provides the motivation for this work. Instead of applying a deterministic control approach based on fluid models, which is prevalent in the existing literature on dissemination control in DTNs [15], [16], [17], [18], [19], we apply a stochastic framework based on diffusion models and the feedback provided by measurements. Moreover, our object of interst is a specific realization of the spreading process rather than the average trajectory described by the fluid model. Note that a specific realization can deviate significantly from the average trajectory (see Figs. 2 and 3), and tracking the specific realization rather than the average trajectory is imperative in such cases.

Motivated by the separation principle, we focus on the tracking of a specific realization of the spreading process by Kalman filtering (which is the first step for LQG stochastic control problems). In general, this tracking problem can be formulated as a Bayesian filtering problem, where the filtering equations could be non-linear. However, by deriving diffusion approximations for the spreading and observation processes, we could formulate it as a Kalman filtering problem where the filtering equations are linear and can be efficiently solved.

Consensus algorithms [34], [35] may also be applied for tracking of the degree of message spreading. However, to be able to track a transient process, as in our case, the averaging must happen at a time scale exponentially faster than the time scale at which the process itself changes [36]. Hence, even though consensus algorithms could be useful in a completely distributed setting, we adopt the Kalman filtering approach in our case, where the observations collected by network nodes are processed by a central entity.

Methods of Spreading and Measurements: We track (over time) the degree of spread of a message in a mobile DTN for the cases of *direct delivery* and *epidemic routing* using Kalman filtering. We assume that there are several *observer* nodes in the network that, upon meeting with *user* nodes, count the number of nodes having the message. Our approach is to keep the counting process *anonymous* (i.e., user identity is not revealed to the observers) and *light-weight* (i.e., the history of meetings with specific users are not maintained) so that a sufficiently large number of (possibly, third-party) observers can be employed for tracking purpose in order to achieve high accuracy.

Mobility Models: The key quantities that characterize the mobility pattern and significantly impact the performance in DTNs are the *inter-meeting times of node pairs* [37]. In [38], [39], the authors provide simulation results to show that, for *random waypoint* mobility models such as the *random direction* and *random walk* mobility models, the inter-meeting times for individual node pairs are well-approximated by i.i.d. exponential random variables.

In [40], [41] and [42], the authors thoroughly examine real-world mobility traces collected in several different network scenarios such as WiFi, vehicular GPS, GSM and Bluetooth. It is concluded that the inter-meeting times for individual node pairs are exponentially distributed [40], [41], but the inter-meeting time aggregated over all node pairs is given by a power law with exponential tail [42]. This dichotomy was recently resolved in [43], where it was shown that *if the inter-meeting times for individual node pairs are exponentially distributed with asymmetric rates and the asymmetric rates are drawn from a Pareto distribution, then the inter-meeting time aggregated over all node pairs is given by a power law with exponential tail*.

Motivated by the above findings in [37], [38], [39], [40], [41], [42] and [43], we consider the following two cases:

- A *homogeneous* mobility model with i.i.d. exponentially distributed pairwise inter-meeting times, and
- A *heterogeneous* mobility model with exponentially distributed pairwise inter-meeting times, where the pairwise rates are drawn from a Pareto distribution.

The homogeneous and heterogeneous mobility models are described in detail in Sec. II-A.

Our Contributions: A major contribution of this paper lies in formulating the tracking of a specific realization of the random message spreading process in DTNs as a Kalman filtering problem. The tracking performance of a Kalman filter, however, is only as good as the equations that are used for representing the system dynamics and the measurements. Often, these two equations are assumed to be known in that the system and measurement matrices, and the auto- and cross-correlation matrices of the process and observation noise sequences are assumed to be known. In reality, assigning appropriate values to those matrices such that they closely represent the actual system and measurements is a non-trivial task. Moreover, using incorrect values make the Kalman filter estimates to deviate significantly from the true realization of the process.

1. *The main contribution of the paper is a rigorous derivation of the* ***system dynamic equation*** *and of the* ***measurement equation*** *for direct delivery and epidemic routing under a homogeneous mobility model (Sec. III, IV and V).*

For the case of homogeneous mobility (see Sec. II-A), we model the process representing the number of nodes having the message as a *density-dependent Markov chain* (see Definition 1 in Appendix A). We also characterize the measurement process as a *doubly stochastic Poisson process* [44]. We derive the diffusion approximations both for the spreading and the observation processes, and obtain the system dynamic and measurement equations by (a) sampling the diffusion approximations at time instants given by a strictly increasing sequence, and (b) obtaining recursive relations. The foregoing method ensures that

2. *The system dynamic and measurement equations derived under homogeneous mobility are closely related to the actual spreading and measurement process, respectively, and the auto- and cross-correlations of the system and measurement noise sequences are accurately determined.*

For the case of heterogeneous mobility (see Sec. II-A), the system dynamic and measurement equations are obtained by using "equivalent symmetric rates" that are expected values of the Pareto distributed heterogeneous rate parameters (see Sec. VI). We show that

3. *The performance of the filter with heterogeneous mobility is comparable to that with homogeneous mobility.*

We evaluate the performance of the filter in terms of the accuracy in detecting certain level-crossing times, i.e., the times when the actual spreading process and the Kalman estimation cross certain levels. This is motivated by the simple control objective where spreading is stopped as soon as a certain fraction (or percentage) of users have received the message. We show that,

4. *In some cases, one may be able to trade off between 'the mean detection delay' and 'the probability of early detection' by varying the number of observers, but, in certain other cases, the number of observers have no significant impact.*

We propose a specific minimization objective by combining detection delay and early detection, and provide several insightful observations (see **O1-O8** in Sec. VII-B).

This paper differs from our earlier work [45] as follows. In [45], the diffusion approximation for the spreading process are obtained by casting it as a 'number in the queue' process. In this paper, we have extended the analysis to include the case of epidemic routing in a unified way by modeling the spreading process as a density-dependent Markov chain. Thus, the analysis (Sec. III-IV) and filtering equations (Sec. V) are significantly generalized and can be applied to a large class of dissemination methods. We also provide a better way of evaluating the tracking performance exhibiting natural trade-offs, and touch upon the issue of choosing the number of observers for optimal performance. Moreover, we also provide a method for tracking under heterogeneous mobility (Sec. VI).

Outline: The remainder of the paper is organized as follows. In Sec. II, we describe our network setting and mobility models. In Sec. III, we provide Markovian characterization of the spreading and observation processes under homogeneous mobility. In Sec. IV, we derive approximations for the spreading and observation processes. In Sec. V, we obtain the discrete time Kalman filter using the diffusion approximations obtained in Sec. IV. The method of equivalent symmetric rates for the case of heterogeneous mobility is described in Sec. VI. Simulation results and discussions with insights are provided in Sec. VII. Sec. VIII concludes the paper. Some notation, definitions and key results from the literature and the proof of a theorem are provided in the appendices.

II. Network Setting and Assumptions

We consider a DTN consisting of S_0 sources, N_0 users, and H_0 observers. The users are mobile. But, the sources and the observers are *static*, and are connected by a high-speed backbone network (see Remark 1 below on mobile sources and observers). One can think of the static sources and observers as base stations (BSs) and/or WiFi access points (APs), and the users as mobile terminals (MTs) and/or vehicles with on-board wireless devices. The number of BSs and/or APs and the number of users often remain constant for a long period of time[2] as compared to the duration of message dissemination, and hence, are taken to be constants. Fig. 1 depicts such a DTN, which represents a sparse deployment of BSs with a sparse population of MTs. We are interested in tracking the dissemination of *non-time-critical* messages such as advertisements and/or informations issued by the local government in public interest.

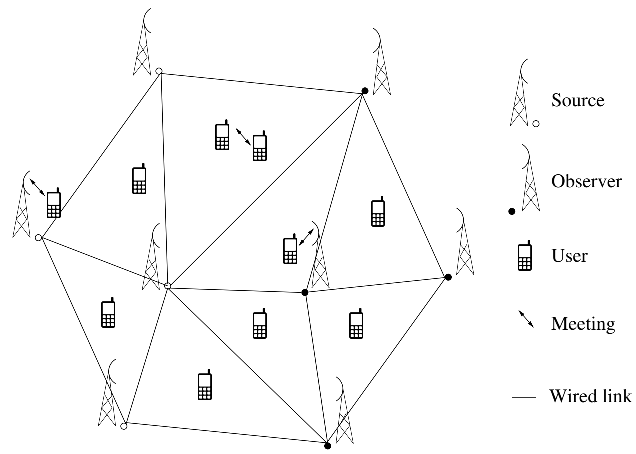

Fig. 1. A DTN consisting of sources, observers and users.

Two nodes are said to "meet" when they come within the communication range of each other (due to the mobility of one or both of the nodes). The mobility model is described in Sec. II-A. We track the degree of spread of *one* given message. At time $t = 0$, the message is only available at the S_0 sources. When a user meets with a source, the user gets a copy of the message with probability p_S (if the user does not already have the message). The probability $1 - p_S$ models the link-layer error for source-user communication. Depending on whether users also spread copies of the message to other users, we study two basic paradigms of message forwarding, namely, *direct delivery* [10] and *epidemic routing* [11]. In direct delivery, only the sources spread the message to the users. In epidemic routing, when a user j, which has the message, meets with another user j', which does not have the message, user j' gets a copy of the message with probability p_U, where $1 - p_U$ models the link-layer error for user-user communication.

Clearly, to be able to control the rate of spreading, the sources need to continuously know the number of users that have already received the message. However, in reality, a source will not be able to know the *exact* number of users that have already received the message. This could be due to one of the following reasons:

[2]The number of BSs and APs typically changes when new BSs/APs are added or old ones removed. Similarly, the number of users changes due to new subscriptions and attrition of existing customers.

(i) The source keeps transmitting a *broadcast* message, and a user receives the message (with probability p_S), but no acknowledgment (ACK) is sent by the user because the message is a broadcast.

(ii) The source transmits a unicast message upon meeting with a user, which receives the message (with probability p_S) and sends an ACK, but the source does not receive the ACK (probably, because the user has already moved out of the range of the source).

(iii) In the case of epidemic routing, the users also spread the message to other users, and the sources cannot count such forwarding events.

Since the sources do not know the exact number of users that have already received the message, they take the help of *external* observers (see Remark 2). At each contact with a user, an observer simply increments a counter if the user is found to have the message. The observers send their observations (counts) to a central entity using the high-speed backbone network. The sum of the observers' counts is used as the measurement.

We assume that *the contact duration of an observer and a user is long enough for the observer to know* ***exactly*** *whether the user has (or does not have) the message*. This is a reasonable assumption since it only requires the exchange of two *short* control packets.

An observer does not (need to) know the identity of a user that it meets with, and does not (need to) keep the history of its meetings with specific users. Hence, one copy at one user might be counted multiple times by the observer(s). However, this *anonymous* and *light-weight* counting scheme allows one to take the help of a large number of third-party observers.

Remark 1. Our analysis methodology would still apply: (i) to mobile sources, if the observers can distinguish between a source and a user, and (ii) to mobile observers, if they can send their counts to the central entity with negligible delay, e.g., by *network-bridging* [46]. ∎

Remark 2. A source can also act as an observer (and vice versa). However, the source-user meetings must be independent of the observer-user meetings.[3] Thus, a node should not act as a source and an observer at the same time, i.e., a node should not give a copy as a source of the message to a user and count the copy as an observer at the same time. A static node can act as a source or an observer for a certain fraction of time. ∎

A. Mobility Models

Motivated by the findings in [38], [39], [37], [40], [41], [42] and [43] (see Sec. I), we consider a *homogenous* and a *heterogeneous* mobility model. In both mobility models, the inter-meeting times of a source-user pair (i,j), an observer-user pair (k,l) and a user-user pair (m,n) are exponentially distributed random variables with means $\beta_{S,i,j}^{-1}$, $\beta_{H,k,l}^{-1}$ and $\beta_{U,m,n}^{-1}$, respectively.

In the homogeneous case, we have $\beta_{S,i,j} = \beta_S$ for every source-user pair (i,j), $\beta_{H,k,l} = \beta_H$ for every observer-user pair (k,l) and $\beta_{U,m,n} = \beta_U$ for every user-user pair (m,n), where β_S, β_H and β_U are constants. The symmetric rate parameters β_S, β_H and β_U can be obtained from real-world traces (see, for example, [47]).

In the heterogeneous case, for every source-user pair (i,j), $\beta_{S,i,j}$ is drawn (or sampled) from the same Pareto distributed random variable $\beta_{S,Pareto}$. Thus, in this case, the inter-meeting times for different source-user pairs are neither independent nor identically distributed. The Cumulative Distribution Function (CDF) of the Pareto distributed rate parameter $\beta_{S,Pareto}$ is given by

$$F_S(u) = P(\beta_{S,Pareto} \le u) = 1 - \left(\frac{x_{min,S}}{u}\right)^{\alpha_S}, \quad (1)$$

where $x_{min,S}$ and α_S are the *scale* and *shape* parameters, respectively. Similarly, for every observer-user pair (k,l), $\beta_{H,k,l}$ is drawn from the same Pareto distributed random variable $\beta_{H,Pareto}$, and for every user-user pair (m,n), $\beta_{U,m,n}$ is drawn from the same Pareto distributed random variable $\beta_{U,Pareto}$, where the CDFs of $\beta_{H,Pareto}$ and $\beta_{U,Pareto}$ are given by

$$F_H(u) = P(\beta_{H,Pareto} \le u) = 1 - \left(\frac{x_{min,H}}{u}\right)^{\alpha_H}, \quad (2)$$

$$F_U(u) = P(\beta_{U,Pareto} \le u) = 1 - \left(\frac{x_{min,U}}{u}\right)^{\alpha_U}. \quad (3)$$

Note that with both mobility models, *the observer-user inter-meeting times are independent of the source-user as well as of the user-user inter-meeting times.*

III. Characterization of the Processes with Homogeneous Mobility

Let $X(t)$ denote the number of *users* that have the message at time t. *Note that $X(t)$ does not include the sources.* Thus, $X(0) = 0$. Let $Y(t)$ denote the *total count of all the observers* up to time t, with $Y(0) := 0$. Henceforth, we shall refer to $\{X(t), t \ge 0\}$ as "the spreading process" and to $\{Y(t), t \ge 0\}$ as "the observation process" or "the measurement process".

From the network and mobility models described in Sec. II, it is clear that, with homogeneous mobility, the process $\{X(t), t \ge 0\}$ is a continuous time Markov chain (CTMC) with state space $\mathcal{S} = \{0, 1, \ldots, N_0\}$ and transition rate matrix $\mathcal{Q} = [q(X, X')]_{X,X' \in \mathcal{S}}$ given by

Direct Delivery:

$$q(X, X+l) = \begin{cases} p_S\beta_S S_0(N_0 - X), & \text{if } l = +1, \\ -p_S\beta_S S_0(N_0 - X), & \text{if } l = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Epidemic Routing:

$$q(X, X+l) = \begin{cases} \left(p_S\beta_S S_0 + p_U\beta_U X\right)(N_0 - X), & \text{if } l = +1, \\ -\left(p_S\beta_S S_0 + p_U\beta_U X\right)(N_0 - X), & \text{if } l = 0, \\ 0, & \text{otherwise.} \end{cases}$$

[3] This is a technical requirement so that Theorem B.2 can be applied.

Given the spreading process $\{X(t)\}$, the observation process $\{Y(t)\}$ is a *doubly-stochastic Poisson process*[4] with (stochastic) intensity function $\beta_H H_0 X(t)$, i.e.,

$$Y(t) = \mathcal{P}_H\left(\beta_H H_0 \int_0^t X(u)du\right), \tag{4}$$

where $\{\mathcal{P}_H(t), t \geq 0\}$ denotes a unit rate Poisson process which is independent of $\{X(t)\}$. The independence can be understood by noting that the process $\{\mathcal{P}_H(\beta_H H_0 t)\}$ represents the total number of observer-user meetings up to time t, which is independent of source-user and user-user meetings that determine $\{X(t)\}$.

IV. Fluid and Diffusion Approximations with Homogeneous Mobility

We obtain the fluid and diffusion approximations for the spreading process $\{X(t)\}$ with homogeneous mobility by invoking limit theorems for *density-dependent Markov chains* [48, Chapter 8], [49, Chapter 11] (see Appendix A for definition and main results). Since the observation process $\{Y(t)\}$ is driven by the spreading process $\{X(t)\}$ (see Eqn. (4)), the fluid and diffusion approximations for $\{Y(t)\}$ are related to that of $\{X(t)\}$ and we obtain them by the *Continuous Mapping Approach* (see Appendix B).

A. Construction of a Family of Density-Dependent Markov Chains Related to the Spreading Process

We construct a family of density-dependent Markov chains related to the spreading process $\{X(t)\}$ as follows. Consider a sequence of networks, indexed by $N = 1, 2, \ldots$, where the N-th network represents a DTN with N mobile users, $S^{(N)}$ static sources and $H^{(N)}$ static observers with homogeneous mobility. For the N-th network, let the source-user, observer-user and user-user inter-meeting times be independent and exponentially distributed random variables with means $1/\beta_S^{(N)}$, $1/\beta_H^{(N)}$, and $1/\beta_U^{(N)}$, respectively. Let, at $t = 0$, the message be only available at the $S^{(N)}$ sources. Let $X^{(N)}(t)$ represent the number of users that have the message at time t, with $X^{(N)}(0) = 0$. Let $Y^{(N)}(t)$ represent the total count of all the observers up to time t, with $Y^{(N)}(0) := 0$. Let the link-layer errors for source-user and user-user communications be modeled by the probabilities $1 - p_S$ and $1 - p_U$, respectively. Clearly, *the network corresponding to $N = N_0$ represents our DTN of interest containing N_0 mobile users as described in Sec. II*. Thus, we have $S^{(N_0)} = S_0$, $H^{(N_0)} = H_0$, $X^{(N_0)}(t) = X(t)$, $Y^{(N_0)}(t) = Y(t)$, $\beta_S^{(N_0)} = \beta_S$, and so on.

Next, we show that, under appropriate conditions, the sequence of processes $\{X^{(N)}(t), t \geq 0\}$, $N = 1, 2, \ldots$, represents a family of density-dependent Markov chains (see Definition 1 in Appendix A).

[4]A *homogeneous* Poisson process over time is characterized by a "constant" or time-independent intensity. A *non-homogeneous* Poisson process is characterized by an intensity function which is a deterministic function of time. A doubly-stochastic Poisson process has an intensity function which is itself a stochastic process [44, Chapter II].

1) Direct Delivery: Consider the sequence of Markov chains $\{X^{(N)}(t)\}$, $N = 1, 2, \ldots$, in the case of direct delivery, where $\{X^{(N)}(t)\}$ has state space $\mathcal{S}^{(N)} = \{0, 1, \ldots, N\}$ and transition rate matrix $\mathcal{Q}^{(N)} = \left[q^{(N)}(X, X')\right]_{X, X' \in \mathcal{S}^{(N)}}$ given by

$$q^{(N)}(X, X+l) = \begin{cases} p_S \beta_S^{(N)} S^{(N)}(N - X), & \text{if } l = +1, \\ -p_S \beta_S^{(N)} S^{(N)}(N - X), & \text{if } l = 0, \\ 0, & \text{otherwise.} \end{cases}$$

The only positive transition rate of $\{X^{(N)}(t)\}$, which is associated with a jump of $+1$, can be rewritten as

$$q^{(N)}(X, X+1) = N\left(p_S \cdot N\beta_S^{(N)} \cdot \frac{S^{(N)}}{N}\right)\left(1 - \frac{X}{N}\right).$$

We stipulate that, for all N,

$$N\beta_S^{(N)} = N_0 \beta_S =: \lambda_S', \quad \text{a constant, and}$$
$$\frac{S^{(N)}}{N} = \frac{S_0}{N_0} =: s_0, \quad \text{a constant, so that}$$
$$p_S \beta_S^{(N)} S^{(N)} = p_S \beta_S S_0 = p_S \lambda_S' s_0 =: \lambda_S, \quad \text{a constant.}$$

Then the transition rate associated with a jump of $+1$ can be written in a density-dependent form as

$$q^{(N)}(X, X+1) \quad = \quad N f_{+1}\left(\frac{X}{N}\right),$$

where

$$f_{+1}(u) = \lambda_S (1 - u). \tag{5}$$

Remark 3. The above scaling amounts to saying that, in the sequence of networks indexed by N, the total rate at which a source meets with all users remains a constant and that the ratio of the number of sources to the number of users remains a constant. This happens, for example, when the number of nodes in the network is increased by increasing the area of the network, but keeping the node density, average node velocity and communication radius for each type of nodes unchanged. ■

2) Epidemic Routing: Consider the sequence of Markov chains $\{X^{(N)}(t)\}$, $N = 1, 2, \ldots$, in the case of epidemic routing, where $\{X^{(N)}(t)\}$ has state space $\mathcal{S}^{(N)} = \{0, 1, \ldots, N\}$ and transition rate matrix $\mathcal{Q}^{(N)} = \left[q^{(N)}(X, X')\right]_{X, X' \in \mathcal{S}^{(N)}}$ given by

$$q^{(N)}(X, X+l) = \begin{cases} \left(p_S \beta_S^{(N)} S^{(N)} + p_U \beta_U^{(N)} X\right)(N - X), & \text{if } l = +1, \\ -\left(p_S \beta_S^{(N)} S^{(N)} + p_U \beta_U^{(N)} X\right)(N - X), & \text{if } l = 0, \\ 0, & \text{otherwise.} \end{cases}$$

The only positive transition rate of $\{X^{(N)}(t)\}$, which is associated with a jump of $+1$, can be rewritten as

$$q^{(N)}(X, X+1) = N\left(p_S \cdot N\beta_S^{(N)} \cdot \frac{S^{(N)}}{N} + p_U \cdot N\beta_U^{(N)} \cdot \frac{X}{N}\right)\left(1 - \frac{X}{N}\right).$$

As before, we stipulate that, for all N,

$$N\beta_S^{(N)} = N_0 \beta_S =: \lambda_S', \quad \text{a constant, and}$$

$$N\beta_U^{(N)} = N_0\beta_U =: \lambda'_U, \quad \text{a constant, and}$$
$$\frac{S^{(N)}}{N} = \frac{S_0}{N_0} =: s_0, \quad \text{a constant, so that}$$
$$p_S\beta_S^{(N)}S^{(N)} = p_S\beta_S S_0 = p_S\lambda'_S s_0 =: \lambda_S, \quad \text{a constant, and}$$
$$p_U N\beta_U^{(N)} = p_U N_0\beta_U = p_U\lambda'_U =: \lambda_U, \quad \text{a constant.}$$

The above scaling can be interpreted similarly as before (see Remark 3). The transition rate associated with a jump of $+1$, can then be written in a density-dependent form as

$$q^{(N)}(X, X+1) \quad = \quad N f_{+1}\left(\frac{X}{N}\right),$$

where

$$f_{+1}(u) = (\lambda_S + \lambda_U u)(1-u). \tag{6}$$

B. Approximations for the Spreading Process

Defining

$$x^{(N)}(t) = \frac{X^{(N)}(t)}{N}, \tag{7}$$

and applying Theorem A.1 in Appendix A, we observe that, as $N \to \infty$, the sequence of processes $\{x^{(N)}(t)\}$, $N = 1, 2, \ldots$, converges, almost surely, to the solution $x(t)$ of the ordinary differential equation (ODE)

$$\frac{dx(t)}{dt} = F_S(x(t)), \quad 0 \le t < \infty, \tag{8}$$

with initial condition $x(0) = 0$, where $F_S(u)$ denotes the drift function corresponding to the spreading process $\{X(t)\}$. Since there is only one positive transition rate of $\{X^{(N)}(t)\}$, which is associated with a jump of $+1$, we have (see (A.1) in Appendix A)

$$F_S(u) = f_{+1}(u), \tag{9}$$

where $f_{+1}(u)$ for direct delivery and epidemic routing are given by (5) and (6), respectively. Solving (8) with initial condition $x(0) = 0$, we obtain $x(t)$ (see Table I).

The deterministic *fluid approximation* for the process $\{X(t)\}$ is given by

$$X(t) = X^{(N_0)}(t) = N_0 x^{(N_0)}(t) \approx N_0 x(t). \tag{10}$$

The fluid approximation $N_0x(t)$ provides a first-order approximation for the spreading process $\{X(t)\}$ and represents its average trajectory. The larger the value of N_0 the better the approximation.

Next, we define

$$v_x^{(N)}(t) = \sqrt{N}\left(x^{(N)}(t) - x(t)\right), \tag{11}$$

where $x^{(N)}(t)$ is given by (7), and $x(t)$ is given in Table I, and observe that $v_x^{(N)}(0) = 0$ for all N since $x^{(N)}(0) = x(0) = 0$. Then, applying Theorem A.2 in Appendix A, we observe that, as $N \to \infty$, the sequence of processes $\{v_x^{(N)}(t)\}$, $N = 1, 2, \ldots$, converges in distribution to the solution $v_x(t)$ of the equation

$$v_x(t) = B_S\left(\int_0^t f_{+1}(x(u))du\right) + \int_0^t F'_S(x(u))v_x(u)du, \tag{12}$$

TABLE I
EXPRESSIONS FOR $x(t)$, $F_S(u)$, $F'_S(u)$, $y(t)$, $\Phi(t)$ AND $\tilde{\Phi}(t)$.

DIRECT DELIVERY

$F_S(u) = \lambda_S(1-u) \Rightarrow F'_S(u) = -\lambda_S$

$x(t) = 1 - e^{-\lambda_S t}, \quad 0 \le t < \infty$

$y(t) = \lambda_H\left(t - \frac{1}{\lambda_S}\left(1 - e^{-\lambda_S t}\right)\right), \quad 0 \le t < \infty$

$\Phi(t) = e^{-\lambda_S t}, \quad 0 \le t < \infty$

$\tilde{\Phi}(t) = \frac{1-e^{-\lambda_S t}}{\lambda_S}, \quad 0 \le t < \infty$

EPIDEMIC ROUTING

$F_S(u) = (\lambda_S + \lambda_U u)(1-u) \Rightarrow F'_S(u) = -\lambda_S + \lambda_U - 2\lambda_U u$

$x(t) = 1 - \frac{\lambda_U + \lambda_S}{\lambda_U + \lambda_S e^{(\lambda_U+\lambda_S)t}}, \quad 0 \le t < \infty$

$y(t) = \frac{\lambda_H}{\lambda_U}\left(\ln\left(\frac{\lambda_U + \lambda_S e^{(\lambda_U+\lambda_S)t}}{\lambda_U+\lambda_S}\right) - \lambda_S t\right), \quad 0 \le t < \infty$

$\Phi(t) = \frac{e^{(\lambda_U+\lambda_S)t}(\lambda_U+\lambda_S)^2}{\left(\lambda_U + \lambda_S e^{(\lambda_U+\lambda_S)t}\right)^2}, \quad 0 \le t < \infty$

$\tilde{\Phi}(t) = \frac{1}{\lambda_S}\left(1 - \frac{\lambda_U+\lambda_S}{\lambda_U + \lambda_S e^{(\lambda_U+\lambda_S)t}}\right), \quad 0 \le t < \infty$

with $v_x(0) = 0$, where $\{B_S(t)\}$ is a standard Brownian motion and $F'_S(u) := \frac{dF_S(u)}{du}$ represents the derivative of the drift function $F_S(u)$ and is given in Table I.

To solve (12), we consider the equivalent (in distribution) stochastic differential equation (SDE)

$$dv_x(t) = F'_S(x(t))v_x(t)dt + \sqrt{f_{+1}(x(t))}dB_S(t), \tag{13}$$

then solve it with initial condition $v_x(0) = 0$, and obtain (see [50, Page 354] for solution method)

$$v_x(t) = \Phi(t)\int_0^t \Phi^{-1}(u)\sqrt{f_{+1}(x(u))}\, dB_S(u), \tag{14}$$

where $\Phi(t)$ is the solution to the ODE

$$\frac{d\Phi(t)}{dt} = F'_S(x(t))\Phi(t), \quad \Phi(0) = 1.$$

The function $\Phi(t)$ is given in Table I.

The *diffusion approximation* for the spreading process $\{X(t)\}$ is given by

$$\begin{aligned} X(t) \quad &= \quad X^{(N_0)}(t) = N_0x(t) + \sqrt{N_0}v_x^{(N_0)}(t) \\ &\approx \quad N_0x(t) + \sqrt{N_0}v_x(t). \end{aligned} \tag{15}$$

The diffusion approximation $N_0x(t) + \sqrt{N_0}v_x(t)$ provides a second-order approximation for the spreading process $\{X(t)\}$ in that the randomness in $\{X(t)\}$ is approximated in terms of a Brownian motion $\{B_S(t)\}$. The larger the value of N_0 the better the approximation.

C. Approximation for the Observation Process

Recall that the Poisson process $\{\mathcal{P}_H(t)\}$ counts the number of observer-user meetings and is given by (4). Since $Y^{(N)}(t)$

denotes the total observer count in the N-th network, we write

$$\begin{aligned} Y^{(N)}(t) &= \mathcal{P}_H\left(\beta_H^{(N)} H^{(N)} \int_0^t X^{(N)}(u)du\right) \\ &= \mathcal{P}_H\left(\lambda_H \int_0^t X^{(N)}(u)du\right), \end{aligned} \tag{16}$$

where we stipulate that, for all $N \geq 1$,

$$N\beta_H^{(N)} = N_0\beta_H =: \lambda'_H, \quad \text{a constant, and}$$
$$\frac{H^{(N)}}{N} = \frac{H_0}{N_0} =: h_0, \quad \text{a constant, so that}$$
$$\beta_H^{(N)} H^{(N)} = \beta_H H_0 = \lambda'_H h_0 =: \lambda_H, \quad \text{a constant.}$$

The above scaling amounts to saying that, in the sequence of networks indexed by N, the total rate at which an observer meets with all users remains a constant and that the ratio of the number of observers to the number of users remains a constant. The scaling can be interpreted similarly as before (see Remark 3).

Next, using the continuous mapping approach (see Appendix B), we obtain an approximation for the observation process $\{Y(t)\}$ in terms of the approximations for the spreading process $\{X(t)\}$. To that end, we define the following:

$$y(t) := \lambda_H \int_0^t x(u)du, \tag{17}$$

$$y^{(N)}(t) := \frac{Y^{(N)}(t)}{N}, \quad \text{and} \tag{18}$$

$$v_y^{(N)}(t) := \sqrt{N}\left(y^{(N)}(t) - y(t)\right). \tag{19}$$

We have the following important result, which we prove in Appendix C.

Theorem 1. *As $N \to \infty$, the sequence of processes $\{v_y^{(N)}(t)\}$, $N = 1, 2, \ldots$, converges in distribution to the process $\{v_y(t)\}$, where*

$$\begin{aligned} v_y(t) &= B_H(y(t)) + \lambda_H \int_0^t v_x(u)du \\ &= \int_0^t \sqrt{\lambda_H x(u)}dB_H(u) + \lambda_H \tilde{\Phi}(t)\Phi^{-1}(t)v_x(t) \\ &\quad -\lambda_H \int_0^t \tilde{\Phi}(u)\Phi^{-1}(u)\sqrt{f_{+1}(x(u))}dB_S(u), \end{aligned} \tag{20}$$

where $\tilde{\Phi}(t) := \int_0^t \Phi(u)du$ and $\{B_H(t)\}$ is a standard Brownian motion that is independent of $\{B_S(t)\}$.

Proof: See Appendix C. ■

The diffusion approximation for the observation process $\{Y(t)\}$ is given by

$$\begin{aligned} Y(t) &= Y^{(N_0)}(t) = N_0 y(t) + \sqrt{N_0} v_y^{(N_0)}(t) \\ &\approx N_0 y(t) + \sqrt{N_0} v_y(t). \end{aligned} \tag{21}$$

Remark 4. Notice that (i) the randomness in the Poisson process $\{\mathcal{P}_H(t)\}$ is approximated in terms of the Brownian motion $\{B_H(t)\}$, and (ii) the diffusion approximation for the observation process $\{Y(t)\}$ depends on the randomness in the spreading process (through $\{B_S(t)\}$) as well as that in the cellection of measurements at observer-user meetings (through $\{B_H(t)\}$). ■

TABLE II
EXPRESSIONS FOR α_k, γ_k, γ'_k, $E[w_k^2]$, $E[w_k n_k]$ AND $E[n_k^2]$ IN CASE OF DIRECT DELIVERY.

$$\alpha_k = e^{-\lambda_S(T_{k+1}-T_k)} \quad ; \quad \gamma_k = \frac{\lambda_H}{\lambda_S}\left(e^{\lambda_S T_k} - 1\right)$$

$$\gamma'_k = \frac{\lambda_H}{\lambda_S}\left(1 - e^{-\lambda_S(T_{k+1}-T_k)}\right)$$

$$E[w_k^2] = e^{-2\lambda_S T_{k+1}}\left(e^{\lambda_S T_{k+1}} - e^{\lambda_S T_k}\right)$$

$$\begin{aligned} E[w_k n_k] = {} & \lambda_H e^{-\lambda_S T_{k+1}}\left(T_{k+1} - T_k\right) \\ & -\frac{\lambda_H}{\lambda_S} e^{\lambda_S T_{k+1}}\left(e^{\lambda T_{k+1}} - e^{-\lambda_S T_k}\right) \end{aligned}$$

$$\begin{aligned} E[n_k^2] = {} & \frac{\lambda_H^2}{\lambda_S^2}\left(e^{\lambda_S T_{k+1}} - e^{\lambda_S T_k}\right) + \lambda_H(1 - \frac{2\lambda_H}{\lambda_S})\left(T_{k+1} - T_k\right) \\ & + \frac{\lambda_H}{\lambda_S}(1 - \frac{\lambda_H}{\lambda_S})\left(e^{-\lambda_S T_{k+1}} - e^{-\lambda_S T_k}\right) \end{aligned}$$

V. DISCRETE TIME KALMAN FILTERING WITH HOMOGENEOUS MOBILITY

In this Section, we derive the discrete time equations for Kalman filtering. To that end, we sample the fluctuation processes $\{v_x(t)\}$ and $\{v_y(t)\}$ at discrete time instants given by a strictly increasing sequence $\{T_k, k \geq 0\}$, with $T_0 = 0$. Any strictly increasing sequence $0 < T_1 < T_2 < \cdots < T_k < \cdots$, may be considered, but the following two possibilities are natural:

1) **Constant Intervals:** In this case, $T_k = kT$, for some constant $T > 0$.
2) **Triggered by Observations:** In this case, the estimates are updated at observer-user meetings.

Defining

$$v_{x,k} := v_x(T_k), \quad k = 0, 1, 2, \ldots\ ,$$

we obtain the *system dynamic equation* from (14) as

$$v_{x,k+1} = \alpha_k v_{x,k} + w_k, \quad k = 0, 1, 2, \ldots, \tag{22}$$

where

$$\alpha_k = \Phi(T_{k+1})\Phi^{-1}(T_k) \tag{23}$$

and

$$w_k = \Phi(T_{k+1}) \int_{T_k}^{T_{k+1}} \Phi^{-1}(u)\sqrt{f_{+1}(x(u))}dB_S(u). \tag{24}$$

Defining

$$v_{y,k} := v_y(T_k), \quad k = 0, 1, 2, \ldots\ ,$$

we obtain from (20) the *measurement equation* as:

$$v_{y,k} = \gamma_k v_{x,k} + z_k, \quad k = 0, 1, 2, \ldots\ , \tag{25}$$

where

$$\gamma_k = \lambda_H \tilde{\Phi}(T_k)\Phi^{-1}(T_k) \tag{26}$$

and $z_k = r_k + s_k$, where

$$r_k = -\lambda_H \int_0^{T_k} \tilde{\Phi}(u)\Phi^{-1}(u)\sqrt{f_{+1}(x(u))}dB_S(u), \quad \text{and}$$

$$s_k = \int_0^{T_k} \sqrt{\lambda_H x(u)}dB_H(u).$$

TABLE III

EXPRESSIONS FOR α_k, γ_k, γ'_k, $E[w_k^2]$, $E[w_k n_k]$ AND $E[n_k^2]$ IN CASE OF EPIDEMIC ROUTING, WHERE $E_k := e^{(\lambda_U+\lambda_S s_0)T_k}$.

$$\alpha_k = \frac{e^{(\lambda_U+\lambda_S)T_{k+1}}(\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_k})^2}{e^{(\lambda_U+\lambda_S)T_k}(\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_{k+1}})^2} \quad ; \quad \gamma_k = \frac{\lambda_H(\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_k})^2(\frac{\lambda_S e^{(\lambda_U+\lambda_S)T_k}-\lambda_S}{\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_k}})}{\lambda_S(\lambda_U+\lambda_S)^2 e^{(\lambda_U+\lambda_S)T_k}}$$

$$\gamma'_k = \frac{\lambda_H(\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_k})((\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_{k+1}})-(\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_k}))}{\lambda_S(\lambda_U+\lambda_S)e^{(\lambda_U+\lambda_S)T_k}(\lambda_U+\lambda_S e^{(\lambda_U+\lambda_S)T_{k+1}})}$$

$$E[w_k^2] = \left[\lambda_S^2(E_{k+1}-E_k)+\lambda_U^2(E_k^{-1}-E_{k+1}^{-1})+2\lambda_U\lambda_S(\lambda_U+\lambda_S)(T_{k+1}-T_k)\right] \times \frac{\lambda_S(\lambda_U+\lambda_S)E_{k+1}^2}{(\lambda_U+\lambda_S E_{k+1})^4}$$

$$E[w_k n_k] = \left[\frac{\lambda_U}{(\lambda_U+\lambda_S)^2}(\frac{\lambda_U}{\lambda_U+\lambda_S}-1)(E_k^{-1}-E_{k+1}^{-1})+\frac{\lambda_S}{\lambda_U+\lambda_S}(\frac{2\lambda_U}{\lambda_U+\lambda_S}-1)(T_{k+1}-T_k)+\frac{(\lambda_S)^2}{(\lambda_U+\lambda_S)^3}(E_{k+1}-E_k)\right] \times \frac{-\lambda_H(\lambda_U+\lambda_S)^2 E_{k+1}}{(\lambda_U+\lambda_S E_{k+1})^2}$$

$$\begin{aligned} E[n_k^2] = {} & \frac{\lambda_H^2}{\lambda_U+\lambda_S}\left[\frac{1}{\lambda_S}(\frac{\lambda_U}{\lambda_U+\lambda_S}-1)^2(E_k^{-1}-E_{k+1}^{-1})+(\frac{2\lambda_U}{\lambda_U+\lambda_S}-2)(T_{k+1}-T_k)+\frac{\lambda_S}{(\lambda_U+\lambda_S)^2}(E_{k+1}-E_k)\right] \\ & +\frac{\lambda_H}{\lambda_U}[log(\lambda_U+\lambda_S E_{k+1})-log(\lambda_U+\lambda_S E_k)]-\frac{\lambda_H}{\lambda_U}\lambda_S(T_{k+1}-T_k) \end{aligned}$$

Defining $n_k := n_{1,k} + n_{2,k}$, where

$$\begin{aligned} n_{1,k} &= -\lambda_H \int_{T_k}^{T_{k+1}} \tilde{\Phi}(u)\Phi^{-1}(u)\sqrt{f_{+1}(x(u))}dB_S(u), \\ n_{2,k} &= \int_{T_k}^{T_{k+1}} \sqrt{\lambda_H x(u)}dB_H(u), \end{aligned}$$

we obtain, $r_{k+1} = r_k + n_{1,k}$, $s_{k+1} = s_k + n_{2,k}$ and

$$z_{k+1} = z_k + n_k. \tag{27}$$

Notice that, for all $i \neq j$, w_i and w_j depend on the value of the Brownian motion $\{B_S(t)\}$ over non-overlapping time intervals. Thus, the process noise sequence $\{w_k\}$ is white. However, the measurement noise sequence $\{z_k\} = \{r_k + s_k\}$ is *sequentially correlated* [51] because the increment sequences $\{n_{1,k}\} = \{r_{k+1} - r_k\}$ and $\{n_{2,k}\} = \{s_{k+1} - s_k\}$ are white. Thus, we adopt *the measurement differencing approach* [51], [32]. Defining

$$v'_{y,k} := v_{y,k} - v_{y,k-1}, \quad \text{for all } k \geq 1, \tag{28}$$

with $v_y(0) = 0$, the *modified measurement equation* becomes

$$\begin{aligned} v'_{y,k+1} &= v_{y,k+1} - v_{y,k} \\ &= \gamma_{k+1}v_{x,k+1} + z_{k+1} - \gamma_k v_{x,k} - z_k \\ &= \gamma_{k+1}(\alpha_k v_{x,k} + w_k) + z_{k+1} - \gamma_k v_{x,k} - z_k \\ &= \gamma'_k v_{x,k} + z'_k \ , \end{aligned} \tag{29}$$

where $\gamma'_k = \gamma_{k+1}\alpha_k - \gamma_k$ and $z'_k = \gamma_{k+1}w_k + n_k$.

The process noise sequence $\{w_k\}$ and the modified measurement noise sequence $\{z'_k\}$ are white with

$$w_k \sim \mathcal{N}(0, Q_k), \quad \text{and} \quad z'_k \sim \mathcal{N}(0, R_k),$$

where $Q_k := E[w_k^2]$, and $R_k := E[(z'_k)^2]$. However, the modified measurement noise is correlated with the process noise. The modified measurement noise z'_l is correlated with the process noise w_k if and only if $l = k$. We define $E[w_k z'_l] := C_k\delta_{kl}$, where δ_{kl} denotes the Kronecker delta function which is equal to 1 if $k = l$, and equal to 0 otherwise. It is easy to see that

$$C_k = E[w_k z'_k] = \gamma_{k+1}E[w_k^2] + E[w_k n_k] \tag{30}$$

and

$$E[(z'_k)^2] = \gamma_{k+1}^2 E[w_k^2] + 2\gamma_{k+1}E[w_k n_k] + E[n_k^2]. \tag{31}$$

The expressions for α_k, γ_k, γ'_k, $E[w_k^2]$, $E[w_k n_k]$ and $E[n_k^2]$ in case of direct delivery and epidemic routing have been summarized in Tables II and III, respectively.

A. The Filtering Algorithm

Let $\hat{v}^+_{x,k}$ denote the Kalman estimate of $v_{x,k}$ at time k after taking into account k measurements up to $v_{y,k}$, i.e.,

$$\hat{v}^+_{x,k} = E[v_{x,k}|v_{y,1}, \ldots, v_{y,k}].$$

Let P^+_k denote the corresponding error covariance, i.e.,

$$P^+_k = E[(v_{x,k} - \hat{v}^+_{x,k})^2].$$

Let

$$\hat{v}_{y,k} := \sqrt{N_0}((\hat{Y}(T_k)/N_0) - y(T_k)),$$

where $\hat{Y}(T_k)$ and $y(T_k)$ denote the actual measurement (i.e., observers' total count) and the value of $y(t)$, respectively, at time T_k. Defining

$$\hat{v}'_{y,k} := \hat{v}_{y,k} - \hat{v}_{y,k-1},$$

we apply Kalman filtering with sequentially correlated measurement noise as follows:

Step 1: Start with $\hat{v}^+_{x,0} = 0$, $P^+_0 = 0$ and $v_y(0) = 0$ (so that $v'_{y,k}$ can be computed for $k \geq 1$ upon the availability of the measurement $v_{y,k}$ for $k \geq 1$).

Step 2: For $k \geq 1$, apply the measurement $\hat{v}'_{y,k}$ to compute $\hat{v}^+_{x,k}$ and P^+_k as follows (see [32, Equation 7.78])

$$\begin{aligned} \hat{v}^+_{x,k} &= \alpha_{k-1}\hat{v}^+_{x,k-1} \\ &\quad + \left[\alpha_{k-1}P^+_{k-1}\gamma'_{k-1} + C_{k-1}\right]\left[(\gamma'_{k-1})^2 P^+_{k-1} + R_{k-1}\right]^{-1} \\ &\qquad \times \left(\hat{v}'_{y,k} - \gamma'_{k-1}\hat{v}^+_{x,k-1}\right) \\ P^+_k &= \alpha_{k-1}^2 P^+_{k-1} + Q_{k-1} \\ &\quad - \left[\alpha_{k-1}P^+_{k-1}\gamma'_{k-1} + C_{k-1}\right]^2\left[(\gamma'_{k-1})^2 P^+_{k-1} + R_{k-1}\right]^{-1} \end{aligned}$$

We obtain the estimates for the process as

$$\hat{X}(T_k) = N_0 x(T_k) + \sqrt{N_0}\hat{v}^+_{x,k},$$

where $\sqrt{N_0}\hat{v}^+_{x,k}$ provides an estimate of the fluctuation of the process about its mean, at time T_k.

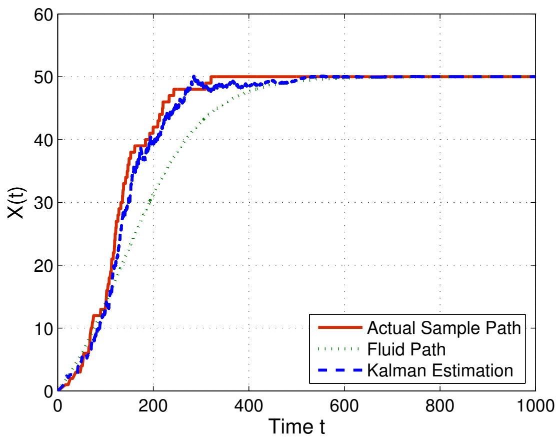

Fig. 2. Value of Kalman filtering as compared to deterministic average fluid path for homogeneous mobility and $S_0 = H_0 = 1$, $N_0 = 50$, $\beta_S = 0.002$, $\beta_U = 0.0002$, $\beta_H = 0.2$.

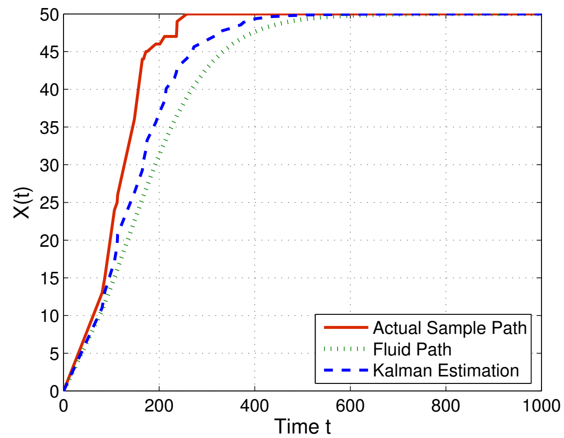

Fig. 3. Value of Kalman filtering as compared to deterministic average fluid path for homogeneous mobility and $S_0 = H_0 = 1$, $N_0 = 50$, $\beta_S = 0.002$, $\beta_U = 0.0002$, $\beta_H = 0.002$.

VI. Heterogeneous Mobility

For the case of heterogeneous mobility, first we obtain "equivalent symmetric rates" for source-user, observer-user and user-user meetings, denoted by $\tilde{\beta}_S$, $\tilde{\beta}_H$ and $\tilde{\beta}_U$, respectively. The equivalent symmetric rates $\tilde{\beta}_S$, $\tilde{\beta}_H$ and $\tilde{\beta}_U$ are, essentially, the expected values of the corresponding Pareto distributed rate parameters $\beta_{S,Pareto}$, $\beta_{H,Pareto}$ and $\beta_{U,Pareto}$ (recall the heterogeneous mobility model described in Sec. II-A). Then, we obtain the system dynamic and measurement equations by applying the framework of Sec. III-V for homogeneous mobility, but using the equivalent symmetric rates $\tilde{\beta}_S$, $\tilde{\beta}_H$ and $\tilde{\beta}_U$.

To evaluate the performance of the filter for the case of heterogeneous mobility, however, we schedule the actual node meetings in simulations (see Sec. VII) according to the heterogeneous mobility model described in Sec. II-A. Thus, with heterogeneous mobility, the system dynamic and measurement equations are **not** as accurate representations of the actual system and mesurements as that with homogeneous mobility, and the performance of the filter may be expected to degrade. However, we show that the performance of the filter for the case of heterogeneous mobility is comparable to that for the case of homogeneous mobility (see Sec. VII).

VII. Performance of the Filter

We developed a customized simulator in MATLAB to simulate the DTN setting described in Sec. II. The sum count of all the observers is used as the measurement and is continuously fed to a central entity where Kalman filtering is performed (recall our assumption in Sec. II that the sources and the observers are static and are connected by a high-speed backbone network). The Kalman filter tracks the process $\{X(t)\}$ triggered by observations.

First we provide the rationale behind tracking a specific realization by using observations instead of just using the detrministic average trajectory, which is the fluid path given by $N_0 x(t)$ (see Table I). Fig. 2 compares the actual realization, the deterministic fluid path and the Kalman estimation for homogeneous mobility and $S_0 = H_0 = 1$, $N_0 = 50$, $\beta_S = 0.002$, $\beta_U = 0.0002$, $\beta_H = 0.2$. We observe that the specific realizations can indeed be different from the deterministic average trajectory, and the observations pertaining to a specific realizations provide valuable information for tracking. If the specific realizations are too different from the average trajectory, then the observations must be taken at higher frequency. For example, as shown in Fig. 3, a lower rate of observations with $\beta_H = 0.002$ results in worse tracking performance. However, thanks to the fluid path, the overall shape of the Kalman filter's estimated trajectory is largely maintained as in the true realization.

In summary, the deterministic average trajectory guides the tracking algorithm between two consecutive observation instants and the observation provide corrections about this deterministic average fluid path.

Next, we quantify the performance of our filter. We obtain results with $S_0 = 1$ source, $N_0 = 50$ users and $H_0 = 1, 2, 3, 4, 5, 10, 15, 30, 50$ observers. Our results are averaged over 1000 simulation runs. In the simulator, we schedule the meetings between every pair of nodes according to exponentially distributed inter-meeting times with an appropriate rate. For the case of homogeneous mobility, all source-user meetings occur at the same rate β_S, all user-user meetings occur at the same rate β_U, and all observer-user meetings occur at the same rate β_H, with $\beta_S = \beta_U = \beta_H = 0.002$. For the case of heterogeneous mobility, for every pair of nodes, we first sample the rate parameter from an appropriate Pareto distribution (as described below) and then schedule their meetings according to exponentially distributed inter-meeting times with the sampled rate.

To be able to compare the performance of the filter for homogeneous and heterogeneous mobilities, we choose the *scale* parameters of the Pareto distributed rate parameters for the heterogeneous case in such a way that their expected values are equal to the corresponding symmetric rates for the homogeneous case. Thus, we set

$$\frac{\alpha_U}{\alpha_U - 1} x_{min,U} = \beta_U, \qquad (32)$$

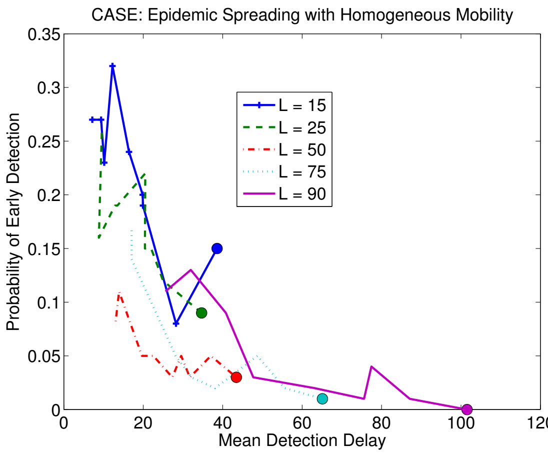

Fig. 4. Trade-off between mean detection delay and probability of early detection for epidemic routing with homogeneous mobility.

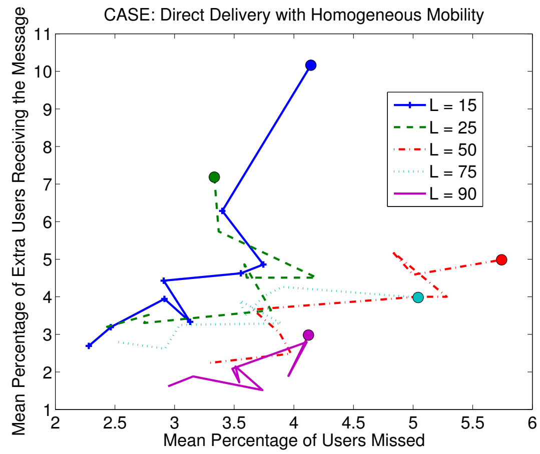

Fig. 5. Mean percentage of "extra users receiving the message" $e_L^{H_0}$ and "missed users" $m_L^{H_0}$ for direct delivery with homogeneous mobility.

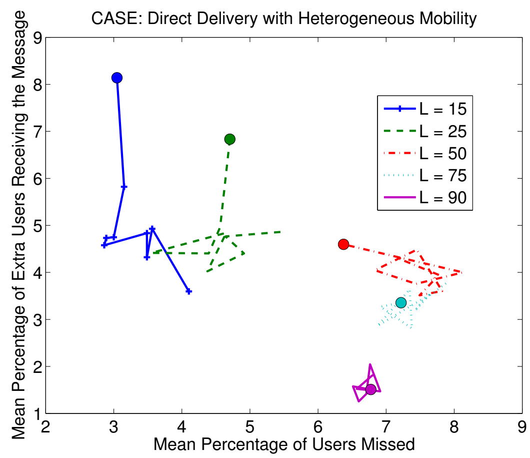

Fig. 6. Mean percentage of "extra users receiving the message" $e_L^{H_0}$ and "missed users" $m_L^{H_0}$ for direct delivery with heterogeneous mobility.

where $x_{min,U}$ and α_U denote the scale and shape parameters of $\beta_{U,Pareto}$ (see Eqn. (3)), and similarly for source-user and obser-user meeting rates. *We make sure that the Pareto distributions have finite variances, which requires that the shape parameter of the Pareto distributions be larger than* 2. *In particular, we take* $\alpha_S = \alpha_H = \alpha_U = 2.5$.

A. Performance Metrics

We evaluate the performance of the Kalman filter in terms of the accuracy in detecting certain level-crossing times, i.e., the times when the actual spreading process and the Kalman estimation cross certain levels. This is motivated by the simple control objective where spreading is stopped as soon as a certain fraction (or percentage) of users have received the message. We consider the target levels $L = 15, 25, 50, 75, 90$.

Let T_L denote the time when $L\%$ of the users actually receive the message (in the simulation). Let $\hat{T}_L^{H_0}$ denote the time when the Kalman filter concludes that $L\%$ of the nodes have received the message, where the superscript H_0 emphasizes the dependence on the number of observers, H_0. Note that *the Kalman estimates can cross a level from below and from above multiple times, but we note the time when the level is crossed from below for the first time*. We study the performance of the filter in terms of the following performance measures:

Late Detection. If $\hat{T}_L^{H_0} > T_L$, then the Kalman filter detects the level-crossing later than when it actually happens. The *mean detection delay*, $D_L^{H_0}$, is defined by

$$D_L^{H_0} = E\left[\hat{T}_L^{H_0} - T_L \mid \hat{T}_L^{H_0} > T_L\right].$$

Also, due to late detection, more than $L\%$ of users receive the message by time $\hat{T}_L^{H_0}$. Then, the *mean percentage of extra users that receive the message* up to Kalman detection time $\hat{T}_L^{H_0}$ is defined by

$$e_L^{H_0} = E\left[X(\hat{T}_L^{H_0}) \times (100/N_0) - L \mid \hat{T}_L^{H_0} > T_L\right],$$

where $X(\hat{T}_L^{H_0})$ denotes the number of users that actually receive the message up to Kalman detection time $\hat{T}_L^{H_0}$. We obtain $D_L^{H_0}$ and $e_L^{H_0}$ by averaging over all those simulation runs in which $\hat{T}_L^{H_0} > T_L$ holds.

Early Detection. If $\hat{T}_L^{H_0} < T_L$, then the Kalman filter wrongly concludes a level-crossing before it actually happens. Then, the *probability of early detection*, $p_{ED}^{H_0}$, is defined by

$$p_{ED,L}^{H_0} = P(\hat{T}_L^{H_0} < T_L).$$

We obtain $p_{ED,L}^{H_0}$ as the fraction of simulation runs in which $\hat{T}_L^{H_0} < T_L$ holds. Also, due to early detection, less than $L\%$ of users have actually received the message by time $\hat{T}_L^{H_0}$. Then, the *mean percentage of users that are missed* due to the early detection is defined by

$$m_L^{H_0} = E\left[L - X(\hat{T}_L^{H_0}) \times (100/N_0) \mid \hat{T}_L^{H_0} < T_L\right],$$

We obtain $m_L^{H_0}$ by averaging over all those simulation runs in which $\hat{T}_L^{H_0} < T_L$ holds.

Remark 5. The case $\hat{T}_L^{H_0} = T_L$ occurs with zero probability, and hence, is ignored. ■

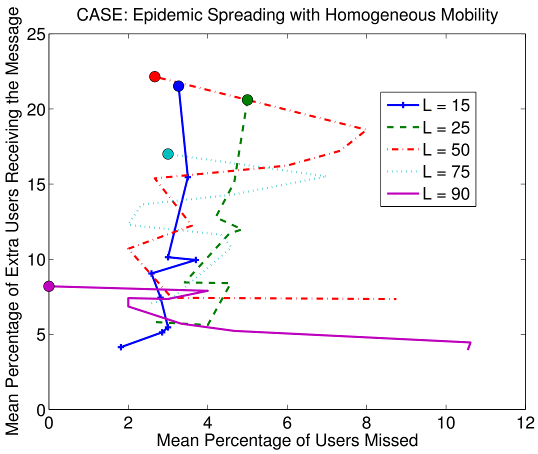

Fig. 7. Mean percentage of "extra users receiving the message" $e_L^{H_0}$ and "missed users" $m_L^{H_0}$ for epidemic routing with homogeneous mobility.

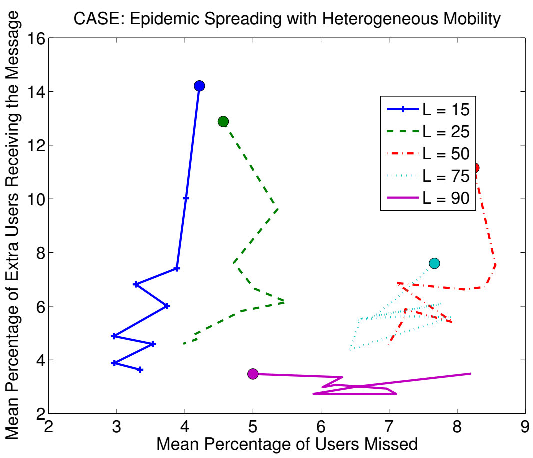

Fig. 8. Mean percentage of "extra users receiving the message" $e_L^{H_0}$ and "missed users" $m_L^{H_0}$ for epidemic routing with heterogeneous mobility.

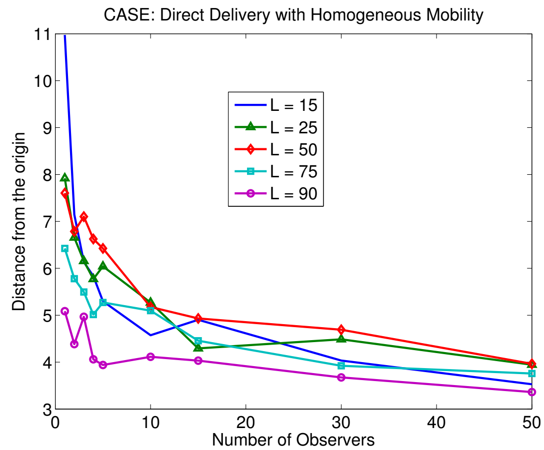

Fig. 9. Distance from the origin for direct delivery with homogeneous mobility.

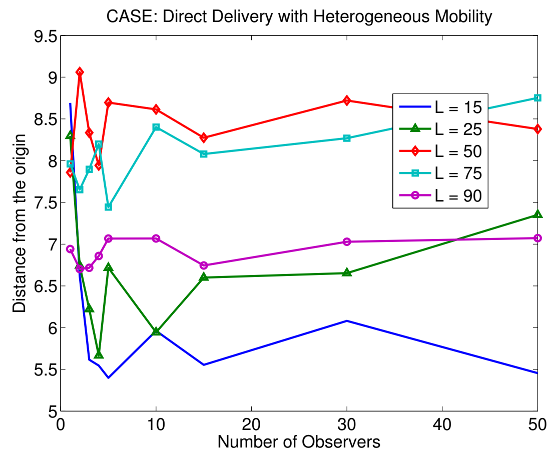

Fig. 10. Distance from the origin for direct delivery with heterogeneous mobility.

B. Results and Observations

Figure 4 depicts the trade-off between the mean detection delay and the probability of early detection for epidemic routing with homogeneous mobility. In Figure 4, results for different target levels have been plotted in different colors and line styles. For each level L, the corresponding polyline starts at the point $(D_L^{H_0}, p_{ED,L}^{H_0})$ corresponding to $H_0 = 1$. The start points are marked with small circles. Then, for each level L, the corresponding polyline connects to the points $(D_L^{H_0}, p_{ED,L}^{H_0})$ corresponding to H_0 = 2, 3, 4, 5, 10, 15, 30, 50, *in that order*. Referring Figure 4, we observe that

O1. There exists a trade-off between the mean detection delay and the probability of early detection. The trade-off is that an increase in the number of observers results in a decrease in the mean detection delay, but the probability of early detection also increases.

However, the trade-off exhibited in Figure 4 may not always exist. For example, we have observed that

O2. With heterogeneous mobility, the number of observers has no significance for higher target levels L = 75 and 90, and less significance for $L = 50$ (plots not shown here due to space constraints, but Figures 6 and 8 shown later provide some evidence to this fact).

If the trade-off exists, one might be able to choose an appropriate number of observers, H_0, to minimize any *weighted combination* of $D_L^{H_0}$ and $p_{ED,L}^{H_0}$. Next, we consider a specific way of combining detection delay and early detection. In Figures 5-8, we show how the number of observers affect the percentage of "extra" and "missed" users. In these plots, we say that *a point* $(e_L^{H_0}, m_L^{H_0})$ *on the polyline for level* L *corresponds to an optimal number of observers if its euclidean distance from the origin is less than (or equal to) that of all other points on that polyline*. In Figures 9-12, we show how the euclidean distance from the origin, $\sqrt{(e_L^{H_0})^2 + (m_L^{H_0})^2}$, varies with the number of observers. Referring to Figures 5-12, we make the following observations:

O3. The maximum mean percentage of missed users, $m_L^{H_0}$, lies between 6-12%. The maximum mean percentage of extra users receiving the message, $e_L^{H_0}$, lies between 9-25%. This worst-case performance can be improved by choosing an appropriate number of observers.

O4. For direct delivery with homogeneous mobility, increasing

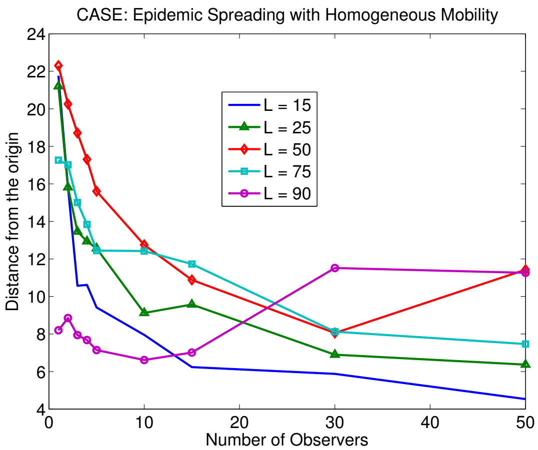

Fig. 11. Distance from the origin for epidemic routing with homogeneous mobility.

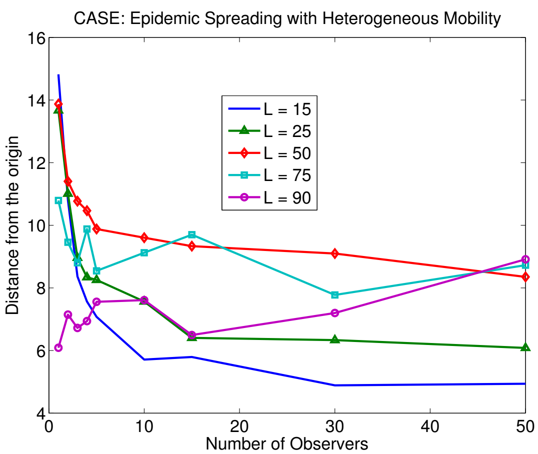

Fig. 12. Distance from the origin for epidemic routing with heterogeneous mobility.

the number of observers results in better performance. In practice, one may choose $H_0 = 10$ observers because the marginal improvement above and beyond $H_0 = 10$ is negligible.

O5. For direct delivery with heterogeneous mobility, increasing the number of observers results in better performance only up to $H_0 = 5$ for $L = 15$ (resp. $H_0 = 4$ for $L = 25$). In other cases, choosing an appropriate value of H_0 is difficult due to highly non-monotonic behaviour of the "distance" metric with H_0. In practice, it is desirable to use the minimum number of observers, i.e., $H_0 = 1$.

O6. For epidemic routing, $H_0 = 30$ appears to be a good level-independent choice except for the level $L = 90$. For $L = 90$, using more than $H_0 = 15$ observers could, in fact, be detrimental.

O7. The appropriate number of observers for epidemic routing, in general, is larger than that for direct delivery because the rate of spreading for epidemic routing is faster than for direct delivery due to spreading by users.

O8. With heterogeneous mobility, the percentage of missed nodes, $m_L^{H_0}$, at a high target level L = 75 or 90, and with large H_0, is higher than that with homogeneous mobility. Indeed, with heterogeneous mobility, some node pairs have much larger (resp. smaller) meeting rates than the average rate. Attainment of a high target level is delayed due to the slower paths, but observations become biased due to the faster paths. The heterogeneity of observer-user meeting rates increases with H_0. This explains the higher value of $m_L^{H_0}$ at higher target levels and with larger H_0.

VIII. Conclusion

In this paper, we have tackled the problem of tracking the degree of message spread in DTNs under direct delivery and epidemic routing. In addition to providing a solid analytical framework, we also provided insightful observations validated with simulations. Several processes of interest in computer and communication networks can be modeled as density-dependent Markov chains. Our most important contribution is that we provide a framework for designing Kalman filters to track such processes.

We have accounted for the delay and noise in collecting measurements by individual observers, but ignored the potential delay and noise in combining the measurements from all the observers by assuming the existence of a high-speed backbone. Extension to the case in which combining the measurements from mobile observers is also governed by mobility can possibly be tackled by applying the continuous mapping approach once more.

References

[1] K. Fall, "A delay-tolerant network architecture for challenged internets," in *SIGCOMM '03*, 2003, pp. 27–34.

[2] "The SARAH project," documents available for download from `http://www-valoria.univ-ubs.fr/SARAH`.

[3] P. Pereira, A. Casaca, J. Rodrigues, V. Soares, J. Triay, and C. Cervelló-Pastor, "From delay-tolerant networks to vehicular delay-tolerant networks," *IEEE Communications Surveys & Tutorials*, no. 99, pp. 1–17, 2011.

[4] C. Caini, H. Cruickshank, S. Farrell, and M. Marchese, "Delay-and disruption-tolerant networking (DTN): An alternative solution for future satellite networking applications," *Proceedings of the IEEE*, no. 99, pp. 1–18, 2011.

[5] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and human mobility in conference environments," in *ACM SIGCOMM workshop on Delay-tolerant networking*, 2005, pp. 244–251.

[6] T. Jonson, J. Pezeshki, V. Chao, K. Smith, and J. Fazio, "Application of delay tolerant networking (DTN) in airborne networks," in *MILCOM 2008*, 2008, pp. 1–7.

[7] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, "Applications, architectures, and protocol design issues for mobile social networks: A survey," *Proceedings of the IEEE*, vol. 99, no. 12, pp. 2130–2158, 2011.

[8] T. Small and Z. Haas, "The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way)," in *MOBIHOC'03*, 2003, pp. 233–244.

[9] A. Pentland, R. Fletcher, and A. Hasson, "Daknet: rethinking connectivity in developing nations," *Computer*, vol. 37, no. 1, pp. 78–83, 2004.

[10] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions On Networking*, vol. 10, no. 4, pp. 477–486, 2002.

[11] A. Vahdat and D. Becker, "Epidemic Routing for Partially-Connected Ad Hoc Networks," in *Duke University, Tech. Rep. CS-200006*, 2000.

[12] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic Routing in Intermittently Connected Networks," in *ACM Mobicom*, 2003.

[13] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Efficient routing in intermittently connected mobile networks: The single-copy case," *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 63–76, 2008.

[14] ——, "Efficient routing in intermittently connected mobile networks: the multiple-copy case," *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 77–90, 2008.
[15] E. Altman, T. Basar, and F. D. Pellegrini, "Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks," *Performance Evaluation*, vol. 67, no. 4, pp. 299–317, 2010.
[16] E. Altman, A. Azad, T. Basar, and F. D. Pellegrini, "Optimal activation and transmission control in delay tolerant networks," in *Proceedings of IEEE INFOCOM*, 2010, pp. 1–5.
[17] A. Karnik and P. Dayama, "Optimal control of information epidemics," in *COMSNETS*, 2012, pp. 1–7.
[18] M. Khouzani, S. Sarkar, and E. Altman, "Optimal control of epidemic evolution," in *2011 Proceedings IEEE INFOCOM*, 2011, pp. 1683–1691.
[19] P. Chen and K. Chen, "Optimal control of epidemic information dissemination in mobile ad hoc networks," in *IEEE GLOBECOM 2011*, 2011, pp. 1–5.
[20] M. Khouzani, S. Sarkar, and E. Altman, "Dispatch then stop: Optimal dissemination of security patches in mobile wireless networks," in *CDC*, 2010, pp. 2354–2359.
[21] M. Khouzani, E. Altman, and S. Sarkar, "Optimal quarantining of wireless malware through reception gain control," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 49–61, 2012.
[22] M. Khouzani, S. Sarkar, and E. Altman, "Optimal dissemination of security patches in mobile wireless networks," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4714–4732, 2012.
[23] W. Kang, F. Kelly, N. Lee, and R. Williams, "Fluid and brownian approximations for an internet congestion control model," in *CDC*, vol. 4, 2004, pp. 3938–3943.
[24] M. Benaim and J. L. Boudec, "A class of mean field interaction models for computer and communication systems," *Performance Evaluation*, vol. 65, no. 11, pp. 823–838, 2008.
[25] R. Darling and J. Norris, "Differential equation approximations for markov chains," *Probability surveys*, vol. 5, pp. 37–79, 2008.
[26] Y. Ko and N. Gautam, "Transient analysis of queues for peer-based multimedia content delivery," *IIE Transactions*, vol. 42, no. 12, pp. 881–896, 2010.
[27] G. Carofiglio, R. Gaeta, M. Garetto, P. Giaccone, E. Leonardi, and M. Sereno, "A fluid-diffusive approach for modelling p2p systems," in *MASCOTS 2006*, 2006, pp. 156–166.
[28] G. Kesidis, T. Konstantopoulos, and P. Sousi, "A stochastic epidemiological model and a deterministic limit for bittorrent-like peer-to-peer file-sharing networks," *Network Control and Optimization (NETCOOP'09)*, pp. 26–36, 2009.
[29] D. Qiu and R. Srikant, "Modeling and performance analysis of bittorrent-like peer-to-peer networks," in *SIGCOMM '04*, 2004, pp. 367–378.
[30] A. Ferragut and F. Paganini, "Content dynamics in p2p networks from queueing and fluid perspectives," in *International Teletraffic Congress (ITC))*, 2012, pp. 1–8.
[31] S. Alouf, E. Altman, C. Barakat, and P. Nain, "Optimal estimation of multicast membership," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2165–2176, 2003.
[32] A. Jazwinski, *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering. Academic Press, 1970, vol. 63.
[33] S. Sethi and G. Thompson, *Optimal control theory: applications to management science and economics*. Springer, 2005.
[34] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2508–2530, 2006.
[35] Y. M. Ko and N. Gautam, "Epidemic-based information dissemination in wireless mobile sensor networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 18, no. 6, pp. 1738–1751, 2010.
[36] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *Automatic Control, IEEE Transactions on*, vol. 54, no. 1, pp. 48–61, 2009.
[37] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance Modeling of Epidemic Routing," *Computer Networks*, vol. 51, pp. 2867–2891, 2007.
[38] R. Groenevelt, P. Nain, and G. Koole, "The Message Delay in Mobile Ad Hoc Networks," *Performance Evaluation*, vol. 62, pp. 210–228, 2005.
[39] M. Ibrahim, A. A. Hanbali, and P. Nain, "Delay and Resource Analysis in MANETs in Presence of Throwboxes," *Performance Evaluation*, vol. 24, no. 9–12, pp. 933–945, October 2007.
[40] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*. ACM, 2009, pp. 299–308.
[41] H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L. M. Ni, "Recognizing exponential inter-contact time in VANETs," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–5.
[42] T. Karagiannis, J.-Y. Le Boudec, and M. Vojnović, "Power law and exponential decay of inter contact times between mobile devices," in *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*. ACM, 2007, pp. 183–194.
[43] A. Passarella and M. Conti, "Analysis of individual pair and aggregate intercontact times in heterogeneous opportunistic networks," *Mobile Computing, IEEE Transactions on*, vol. 12, no. 12, pp. 2483–2495, 2013.
[44] P. Bremaud, *Point Processes and Queues: Martingale Dynamics*. Springer, 1981.
[45] A. Ali, E. Altman, T. Chahed, D. Fiems, M. Panda, and L. Sassatelli, "Estimating file-spread in delay tolerant networks under two-hop routing," in *NETWORKING 2012*. Springer, 2012, pp. 277–290.
[46] K. Harras and K. Almeroth, "Transport layer issues in delay tolerant mobile networks," *Lecture Notes in Computer Science (3976), NETWORKING 2006*, pp. 463–475, 2006.
[47] E. Hernández-Orallo, J.-C. Cano, C. T. Calafate, and P. Manzoni, "A representative and accurate characterization of inter-contact times in mobile opportunistic networks," in *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*. ACM, 2013, pp. 309–316.
[48] T. G. Kurtz, *Approximation of Population Processes*. Society for Industrial and Applied Mathematics, 1981.
[49] S. Ethier and T. Kurtz, *Markov Processes: Characterization And Convergence*, ser. Wiley Series in Probability And Statistics. Wiley Interscience, 2005.
[50] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 1988.
[51] A. E. Bryson and L. J. Henrikson, "Estimation using sampled data containing sequentially correlated noise," *Journal of Spacecraft and Rockets*, vol. 5, pp. 662–665, 1968.
[52] W. Whitt, *Stochastic-Process Limits*. Springer, 2002.

APPENDIX A
DENSITY-DEPENDENT MARKOV CHAINS

In this appendix, we recall the definition of and two limit theorems for the so-called *density dependent Markov chains* [48], [49]. First, we fix some notation. The set of integers (resp. real numbers) is denoted by $\mathbb{Z}$ (resp. $\mathbb{R}$). The space of d-dimensional vectors with integer (resp. real) components is denoted by $\mathbb{Z}^d$ (resp. $\mathbb{R}^d$). The absolute value of a scalar b is denoted by $|b|$. The euclidean norm of a vector z is denoted by $\|z\|$. The transpose of a vector z (resp. a matrix G) is denoted by z^T (resp. G^T).

Consider a one-parameter family of continuous time Markov chains $\{Z^{(n)}(t), t \geq 0\}$, indexed by $n = 1, 2, \ldots$, where $\{Z^{(n)}(t)\}$ has state space $\mathcal{S}^{(n)} \subset \mathbb{Z}^d$ and transition rate matrix $\mathcal{Q}^{(n)} = [q^{(n)}(Z, Z')]_{Z,Z' \in \mathcal{S}^{(n)}}$.

Definition 1 ([48], [49])**.** The family of Markov chains $\{Z^{(n)}(t), t \geq 0\}$, $n = 1, 2, \ldots$, is called density-dependent if there exist a subset $\mathcal{R}$ of $\mathbb{R}^d$ and continuous functions f_l, $l \in \mathbb{Z}^d$, with $f_l : \mathcal{R} \to \mathbb{R}$, such that

$$q^{(n)}(Z, Z + l) = n f_l \left(\frac{Z}{n} \right), \qquad l \neq 0. \qquad \square$$

In practice, instead of considering all possible $l \in \mathbb{Z}^d$, one only needs to consider a much smaller set $\mathcal{L}$, where

$$\mathcal{L} = \{l \in \mathbb{Z}^d : l \neq 0, q^{(n)}(Z, Z + l) \neq 0 \text{ for some } Z \in \mathcal{S}^{(n)}\},$$

that is, one considers only those values of $l \in \mathbb{Z}^d$ that correspond to the transitions that can actually occur in the Markov chain with positive rate. For all $l \notin \mathcal{L}$, one can set

f_l identically equal to zero. Henceforth, we only consider transitions with positive rates.

Define the *drift function* $F(\cdot)$ by

$$F(u) := \sum_l l f_l(u), \quad u \in \mathcal{R}. \tag{A.1}$$

Note that $F(u) = (F_1(u), \ldots, F_d(u))^T$ is a d-dimensional column vector of functions, because l is a d-dimensional column vector. Define the *Jacobian matrix* of F by

$$J_F(u) := (\nabla F_1(u), \ldots, \nabla F_d(u))^T, \quad u \in \mathcal{R}, \tag{A.2}$$

where $\nabla F_i(u) = \left(\frac{\partial F_i(u)}{\partial u_1}, \ldots, \frac{\partial F_i(u)}{\partial u_d}\right)$ denotes the *gradient* of the function F_i, and $u = (u_1, \ldots, u_d)$.

Defining the *density process* $\{z^{(n)}(t)\}$ by $z^{(n)}(t) := \frac{Z^{(n)}(t)}{n}$, we recall the Functional Strong Law of Large Numbers (FSLLN) for density-dependent Markov chains (see [49, Chapter 11, Theorem 2.1]).

Theorem A.1 (Ethier & Kurtz [49]). *Suppose that for each compact set $K \subset \mathcal{R}$,*

$$\sum_l \|l\| \sup_{u \in K} f_l(u) < \infty,$$

and there exists $M_K > 0$ such that

$$\|F(u) - F(u')\| \le M_K \|u - u'\|, \quad \forall u, u' \in K.$$

Suppose also that

$$\lim_{n \to \infty} z^{(n)}(0) = z_0,$$

and $z(\cdot)$ satisfies

$$z(t) = z_0 + \int_0^t F(z(u))du, \quad t \ge 0. \tag{A.3}$$

Then, for every t, $0 \le t < \infty$,

$$\lim_{n \to \infty} \sup_{0 \le s \le t} \|z^{(n)}(s) - z(s)\| = 0,$$

almost surely. □

Remark 6. Theorem A.1 says that, when the drift function $F(\cdot)$ is uniformly bounded and Lipschitz continuous over compact subsets of $\mathcal{R}$, then, as $n \to \infty$, the density process $\{z^{(n)}\}$ converges uniformly over compact subsets to a deterministic function $z(\cdot)$, almost surely, to a process $\{z(t)\}$ given by (A.3). ■

Defining the *deviation process* $\{v_z^{(n)}(\cdot)\}$ by

$$v_z^{(n)}(t) = \sqrt{n}(z^{(n)}(t) - z(t)),$$

where $\{z^{(n)}(t)\}$ is the density process and $\{z(t)\}$ is its deterministic limit given by (A.3), we recall the Functional Central Limit Theorem (FCLT) for density-dependent Markov chains (see [49, Chapter 11, Theorem 2.3]).

Theorem A.2 (Ethier & Kurtz [49]). *Suppose that for each compact set $K \subset \mathcal{R}$,*

$$\sum_l \|l\|^2 \sup_{w \in K} f_l(u) < \infty,$$

and that the f_l, $l \in \mathcal{L}$, and J_F are continuous. Suppose also that

$$\lim_{n \to \infty} v_z^{(n)}(0) = v_{z_0},$$

where v_{z_0} is a constant. Then, as $n \to \infty$, $\{v_z^{(n)}(\cdot)\}$ converges in distribution to $v_z(\cdot)$, where $v_z(\cdot)$ is the solution to the stochastic integral equation

$$\begin{aligned} v_z(t) \quad = \quad & v_{z_0} + \sum_l l B_l \left(\int_0^t f_l(z(u))du \right) \\ & + \int_0^t J_F(z(u)) v_z(u) du, \end{aligned} \tag{A.4}$$

where $\{B_l(\cdot)\}$, $l \in \mathcal{L}$, are independent standard Brownian motions (each corresponding to a transition with positive rate) and $z(t)$ is given by (A.3). □

Appendix B
Continuous Mapping Approach

The continuous mapping approach exploits previously established stochastic-process limits and the Continuous Mapping Theorem to obtain new stochastic-process limits [52]. We recall a particular version of the Continuous Mapping Theorem (CMT) which specifies the conditions under which convergence is preserved for composition plus addition. First, we define some notation.

The identity mapping is denoted by ι. The empty set is denoted by $\emptyset$. The composition of two functions f and g is denoted by $f \circ g$. The set of discontinuities of a function g is denoted by $Disc(g)$. Let (E, m) denote a metric space with metric m. The space of functions $\phi : [0, \infty) \to E$ that are continuous are denoted by $\mathbb{C}_E[0, \infty)$. The space of functions $\phi : [0, \infty) \to E$ that are right-continuous and have left limits are denoted by $\mathbb{D}_E[0, \infty)$. We are particularly interested in the case $E = \mathbb{R}^d$, and use the simple notation $\mathbb{C}^d$ (resp. $\mathbb{D}^d$) to denote $\mathbb{C}_{\mathbb{R}^d}[0, \infty)$ (resp. $\mathbb{D}_{\mathbb{R}^d}[0, \infty)$). The subset of functions in $\mathbb{C}^1$ that are nondecreasing (resp. strictly increasing) is denoted by $\mathbb{C}^1_{\uparrow}$ (resp. $\mathbb{C}^1_{\uparrow\uparrow}$). The subset of functions in $\mathbb{D}^1$ that are nondecreasing (resp. strictly increasing) is denoted by $\mathbb{D}^1_{\uparrow}$ (resp. $\mathbb{D}^1_{\uparrow\uparrow}$). The cartesian product of two spaces S and S' is denoted by $S \times S'$. Convergence in distribution is denoted by $\Rightarrow$. A stochastic process can be viewed as a *random element* in a suitable space. The joint convergence in distribution of the sequence of random elements $\{(z_a^{(n)}, z_b^{(n)}, z_c^{(n)})\}$, $n \ge 1$, to a random element $\{(z_a, z_b, z_c)\}$ is denoted by $(z_a^{(n)}, z_b^{(n)}, z_c^{(n)}) \Rightarrow (z_a, z_b, z_c)$.

Next, we recall the particular version of CMT that we shall apply in Appendix C (see [52, Theorem 13.3.1]).

Theorem B.1 (Whitt [52]). *Let ψ_1, ψ_3 and $\psi_1^{(n)}$, $n \ge 1$, be random elements of $\mathbb{D}^d$; let ψ_2, $\psi_2^{(n)}$ and $\psi_3^{(n)}$, $n \ge 1$, be random elements of $\mathbb{D}^1_{\uparrow}$; and let $c_n \in \mathbb{R}^d$ for $n \ge 1$. If*

$$\left(\psi_1^{(n)} - c_n \iota, \psi_2^{(n)}, c_n(\psi_2^{(n)} - \psi_3^{(n)})\right) \Rightarrow (\psi_1, \psi_2, \psi_3)$$

in $\mathbb{D}^d \times \mathbb{D}^1_{\uparrow} \times \mathbb{D}^d$, and $\psi_2 \in \mathbb{C}^1_{\uparrow\uparrow}$ and

$$P\left(Disc(\psi_1 \circ \psi_2) \cap Disc(\psi_3) = \emptyset\right) = 1,$$

then

$$\psi_1^{(n)} \circ \psi_2^{(n)} - c_n \psi_3^{(n)} \Rightarrow \psi_1 \circ \psi_2 + \psi_3 \quad \text{in } \mathbb{D}^d. \qquad \square$$

When applying Theorem B.1 in Appendix C, we shall establish the requirement of joint convergence in distribution by invoking independence (Theorem B.2) and convergence in distribution to a deterministic limit for one of the components of the joint random elements (Theorem B.3). Theorems B.2 and B.3 are from [52, Theorem 11.4.4] and [52, Theorem 11.4.5], respectively.

Theorem B.2 (Whitt [52])**.** *Let, $\forall n \geq 1$, $\psi_1^{(n)}$ and $\psi_2^{(n)}$ be **independent** random elements of separable metric spaces (S', m') and (S'', m''), respectively. Then there is joint convergence in distribution $(\psi_1^{(n)}, \psi_2^{(n)}) \Rightarrow (\psi_1, \psi_2)$ in $S' \times S''$ if and only if $\psi_1^{(n)} \Rightarrow \psi_1$ in S' and $\psi_2^{(n)} \Rightarrow \psi_2$ in S''.* ■

Theorem B.3 (Whitt [52])**.** *Suppose that $\psi_1^{(n)} \Rightarrow \psi_1$ in a separable metric space S' and $\psi_2^{(n)} \Rightarrow \psi_2$ in a separable metric space S'', where ψ_2 is deterministic. Then*

$$\left(\psi_1^{(n)}, \psi_2^{(n)}\right) \Rightarrow (\psi_1, \psi_2). \qquad \square$$

Appendix C
Proof of Theorem 1

We observe that $v_y^{(N)}(t)$ can be rewritten as

$$v_y^{(N)}(t) = (\phi_1^{(N)} \circ \phi_2^{(N)})(t) - c_N y(t),$$

where $\phi_1^{(N)}(t) = \frac{\mathcal{P}_H(Nt)}{\sqrt{N}}$, $\phi_2^{(N)}(t) = \lambda_H \int_0^t x^{(N)}(u)du$, where $x^{(N)}(t)$ is given by (7), and $c_N = \sqrt{N}$.

By the FCLT for Poisson processes, under CLT-type scaling, a sequence of scaled Poisson processes converges to a standard Brownian motion as the scaling parameter goes to infinity. Thus, as $N \to \infty$, we have

$$(\phi_1^{(N)} - c_N \iota)(t) = \frac{\mathcal{P}_H(Nt) - Nt}{\sqrt{N}} \Rightarrow B_H(t),$$

where $\{B_H(t)\}$ is a standard Brownian motion that approximates the randomness in observer-user meetings, and is independent of $\{B_S(t)\}$ (which approximates the randomness in source-user and user-user meetings).

By the continuity of the integration operator, as $N \to \infty$, we have,

$$c_N\left(\phi_2^{(N)} - y\right)(t) = \lambda_H \int_0^t v_x^{(N)}(u)du \Rightarrow \lambda_H \int_0^t v_x(u)du,$$

where $v_x^{(N)}(t)$ and $v_x(t)$ are given by (11) and (12), respectively. Recalling the independence of $\{\mathcal{P}_H(t)\}$ and $\{X(t)\}$, we observe that, for all $N \geq 1$, $\{\phi_1^{(N)}(t)\}$ and $\{\phi_2^{(N)}(t)\}$ are independent processes. Then, applying Theorem B.2 (see Appendix B), we obtain, as $N \to \infty$,

$$\left(\left(\phi_1^{(N)} - c_N \iota\right), c_N\left(\phi_2^{(N)} - y\right)\right) \Rightarrow \left(B_H, \lambda_H \int v_x(u)du\right).$$

Next, we observe that, $x^{(N)}(t) \to x(t)$, almost surely, implies that $x^{(N)}(t) \Rightarrow x(t)$, since almost sure convergence implies convergence in distribution. Then, by continuous mapping, as $N \to \infty$, we have

$$\phi_2^{(N)}(t) \Rightarrow y(t).$$

Observing that $y(t)$ is a deterministic function, we apply Theorem B.3 (see Appendix B), to conclude that

$$\left(\left(\phi_1^{(N)} - c_N \iota\right), \phi_2^{(N)}, c_N\left(\phi_2^{(N)} - y\right)\right) \Rightarrow \left(B_H, y, \lambda_H \int v_x(u)du\right).$$

Finally, observing that $y(t)$ is strictly increasing and continuous, we apply Theorem B.1 (see Appendix B), and obtain (20). ■