Analytic Latency Model for Message Dissemination in Opportunistic Networks

Qi Wang
Key Laboratory of Computer Network and Information
Integration in Southeast University, Ministry of Education
School of Computer Science and Engineering
Southeast University
Nanjing, China
Email: qiwang@seu.edu.cn

Sanfeng Zhang*
Key Laboratory of Computer Network and Information
Integration in Southeast University, Ministry of Education
School of Computer Science and Engineering
Southeast University
Nanjing, China
Email: sfzhang@seu.edu.cn

***Abstract*—In opportunistic networks, messages are delivered based on a store-forward-carry paradigm. Network topology varies with time and dynamism of the links between nodes has a primary effect on the message delay, one of the network performance indicators. To get a precise estimation of the network latency with less complexity, a rigorous framework modeling the information propagation process is developed. The model is based on a sophisticated 2^{N-1}-state Markov chain and yields a closed-form expression under non-homogenous assumption, as well as an asymptotic formula under homogenous assumption. Finally, to assess our model's scalability and reliability, analytical results are validated by simulations on a real-life human mobility trace and two standard mobility models. The results demonstrate that the model predicts the routing performance accurately for networks of different size and mobility models.**

I. Introduction

With the number of mobile devices growing at a high rate, the scheme for message propagation between mobile nodes becomes a sophisticated problem. In mobile ad hoc networks, nodes are commonly intermittently connected and complete end-to-end paths may not exist. To cope with dynamism, Opportunistic Networks (Oppnets) [1] utilise a store-forward-carry paradigm a node can store the received messages in the local buffer, carry messages while moving, then distribute messages to the nodes that it encounters, either relay nodes or destination nodes. In such a way, messages can be passed successfully solely based on the short-range transmission between mobile nodes without any extra infrastructure.

Modeling and analysing such an interesting evolution of dynamic networks are quite challenging due to the various movement patterns and other minor factors, yet instructive and receives much attention. Early, an asymptotic approximation of message delay under the epidemic protocol is obtained in [2][3] through exploiting simple Markov chain and ordinary differential equations (ODEs). But the frameworks proposed in [2][3] can only be applied to describe the ideal mobility model, and the analytical result is accurate when network size $N \to \infty$. Later, some presented the asymptotic bounds on the metrics of the network performance[4][8], such as the information propagation speed. More recently, prediction results given in [9] are close to real performance but only for two idea conditions. [5][6] set up edge-Markovian models with various parameters, which can quantify different parameters' effect on the latency but the results have high calculational complexity and lack practicality. And the model proposed in [7] has the same feature. Hence, our major motivation is obtaining low-complexity expressions of average latency that both work in real and idea conditions.

Analysis on other information spreading processes in dynamic networks also inspired our work. In recent years, time-graphs [13] discretize the continuous-time process and provides a general, real description of dynamic network. The expected hitting time of a random walk in the time-graph is obtained based on a non-homogenous Markov Chain in [10]. Non-homogenous model is more close to the real-world network. The virus spread process in networks of size N was investigated in [11] through an exact 2^{N-1}-state Markov chain.

In this paper, we focus on the general structure of networks and propose a new rigorous stochastic framework, capturing fundamental characteristics of Oppnets. The general framework is applied to different cases and expected message delay in opportunistic networks is calculated based on a 2^{N-1}-state Markov chain. The complexity of our result does not increase with the network size and can be applied to both real and ideal conditions, which is the our novel contribution.

The rest of this paper is structured as follows. Next section introduces the related work on the calculating the average latency in Oppnets. In Section III we will describe our proposed model, then deduce the expected message delay under different situations. In Section IV, a real-life trace and two classical mobility models are applied to assess prediction reliability of our model. Questions remained are discussed as well. Section V concludes the paper.

II. RELATED WORK

Many models have been developed to study the information dissemination process and routing performance in opportunis-

*Corresponding author
This work is supported by the National Natural science Foundation of China under Grant No. 61300200, No.61472080 and No.61370209.

tic networks since epidemic routing was proposed. The average latency for message dissemination is one of the most important indicators. Many analytical results about it have been yield based on different theoretical frameworks. Here we review some basic models and important results.

Markovian models are widely exploited to study the performance of different routing schemes. Authors of [2] considered three standard mobility models and showed that for these three ideal mobility models, meeting process between any two nodes follows a Poisson process at the same intensity λ. Under this assumption, they developed a simple stochastic model based on modeling the number of copies in the network as a Markov chain. The asymptotic approximation expression of average message delay under epidemic protocol is

$$\mathbb{E}(T) = \frac{1}{\lambda N}(\ln(N) + \gamma + \mathrm{O}(\frac{1}{N})). \tag{1}$$

The model gives quite accurate prediction for three standard models.

By deriving ordinary differential equation (ODE) models as limiting processes of Markovian models,[3] obtained close-form expressions regarding packet-delivery delay. The explicit expression of average delivery delay under epidemic protocol given in [3] is $\mathbb{E}(T) = \frac{\ln N}{\lambda(N-1)}$, which matches the formula (1) obtained in [2] as $N \to \infty$.

In recent years, [5] presented a more realistic and complex framework and gave a detailed expression of average information dissemination delay. The model has various parameters, such as link generating speed, message size and so on. But the numerical calculation becomes impractical to some extent when the network size N is large. Similar results obtained in [6], [7].

In addition, [10] exploited time-graph to characterise the dynamic network and derived the expected hitting time for a random walk in an evolving time-graph (following a non-homogenous Markov Chain). In this paper, we aim to build a more scalable model and give a more precise, practical estimation of average message delivery latency.

III. NETWORK MODEL

In this section, we describe a generic model that does not assume nodes are mutually independent and homogeneous. Consider network size is N and the link between two nodes is bidirectional. As stated in [2], the transmission time of a message between two nodes is negligible compared to the time interval of two successive meetings. Hence, it‘s reasonable to assume that in the i-th time snapshot, if and only if the link between two nodes is open, the package can be exchanged successfully. Furthermore, periodic contact patterns are universally observed in different scenarios [12][14]. So another assumption is the periodic condition, which implies if period time is T, then the $(t+T)$-th time snapshot is same as the t-th time snapshot.

A. Definition of Model

Without any loss of generality, assume at starting time 0, node s creates a message whose destination is node d, and its state is infected $X_s = 1$. At the next time instant (after the unit interval Δt), node u meets node s, then node s transfers message to node u. As soon as node u receives the message, its state will change from $X_u = 0$ (healthy) to $X_u = 1$ (becomes infected). Let $\Delta t = 1$. At each time t, a node is in one of the two states. Every infected node can infect other nodes when they encounter. Let Y_t be the state of the whole network at time t. Define $Y_t = i, \quad i = \sum\limits_{k=1}^{N} X_k^t 2^{k-1}$.

Since node s is infected at start time 0, so the minimum value of Y_t and the initial state of the network equals to 2^{s-1}. Hence, the system state space is $E = \left\{2^{s-1},\ 2^{s-1}+1,\ \dots\ ,\ 2^N - 1 - 2^{s-1}\right\}$. Denote the probability of the link between node u and v being open at time t as $p_{u,v}^t (0 < p_{u,v}^t < 1)$. The probability of node u being infected at time t is denoted as p_u^t, and the probability of node u not being infected at time t is ζ_u^t. We find,

$$\begin{aligned} 1 - p_u^t = &\quad \zeta_u^t \\ &= \prod_{v=1; v\neq u}^{N} \left[\left(1 - x_v^{t-1}\right) + x_v^{t-1}\left(1 - p_{u,v}^t\right)\right] \\ &= \prod_{v=1; v\neq u}^{N} \left(1 - x_v^{t-1} p_{u,v}^t\right). \end{aligned}$$

Then the message propagation process can be defined as a discrete-time non-homogeneous Markov chain (NHMC) on the network state space E. Let S_{t+1} be the set of nodes being newly infected at time $t+1$, D_{t+1} be the set of nodes that stay healthy at $t+1$:

$$S_{t+1} = \left\{k : 1 \le k \le N, x_k^t = 0, x_k^{t+1} = 1\right\},$$

$$D_{t+1} = \left\{k : 1 \le k \le N, x_k^t = 0, x_k^{t+1} = 0\right\}.$$

The single step transition probabilities $p_{t,t+1}(i,j) = P(Y_{t+1} = j \mid Y_t = i)$ are:

$$\mathbb{P}_{t,t+1}(i,j) = \begin{cases} \prod\limits_{\substack{k=1;\\ k\in D_{t+1}}}^{N} \zeta_k^{t+1}, & j = i \\ \prod\limits_{\substack{k=1;\\ k\in D_{t+1}}}^{N} p_k^{t+1} \prod\limits_{\substack{k=1;\\ k\in S_{t+1}}}^{N} \zeta_k^{t+1}, & j = i + \sum\limits_{\substack{k=1;\\ k\in S_{t+1}}}^{N} 2^{k-1} \\ 0, & else \end{cases}$$

where $i, j \in E$. States of the network (E) are divided into two subsets:

- Absorbing states ($A \subset E$ and $A \neq \phi$): Destination node d has received the message, namely $Y_t = i, i = \sum\limits_{k=1}^{N} X_k^t 2^{k-1}$ is an absorbing state if $X_d = 1$.
- Transient states ($B = E \backslash A$): Destination node d hasn‘t received the message, is still healthy.

The Markov state diagram of a network with $N = 4$ nodes as an example is shown in Fig. 1. Let ϕ denote the hitting

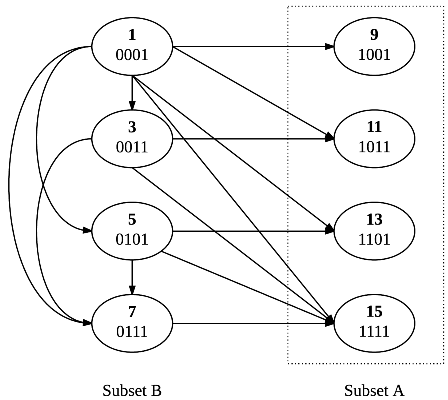

Fig. 1. State transition diagram of a network with 4 nodes as an illustration (using the notation in [11]). All states in the graph are represented in decimal and binary format. Suppose the message is created by node 0 at time 0 and needs be delivered to node 3, so the initial state of network is $Y_0 = 1$. States on the left column are transient states, because the node 3 has not received the message at those states. Similarly, the states on the right column are absorbing states since node 3 has received the message at those states.

time when destination node d receives the message. That is when the chain is first in an absorbing state:

$$\Theta(\omega) = \inf\{t > 0; \quad Y_t(\omega) \in A\}.$$

Adopting the strategy in [15], the transition matrix can be partitioned as

$$\mathbf{p}_t = \begin{pmatrix} \mathbf{p}_t^B & \mathbf{p}_t^{BA} \\ \mathbf{p}_t^{AB} & \mathbf{p}_t^A \end{pmatrix} = \begin{pmatrix} \mathbf{p}_t^B & \mathbf{p}_t^{BA} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

For notational simplicity, define $\mathbf{Q}_t = \mathbf{p}_t^B$,$\mathbf{H}_t = \mathbf{p}_t^{BA}$. Then transition matrix is

$$\mathbf{p}_t = \begin{pmatrix} \mathbf{Q}_t & \mathbf{H}_t \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

B. Non-homogenous Situation

In this part, a detailed expression of calculating message delay in a generic and non-homogeneous situation will be deduced.

Theorem 1: Under epidemic protocol,

$$\mathbb{E}(T) = [1 + (\mathbf{I} - \prod_{i=0}^{i=T-1} \mathbf{Q}_i)^{-1} (\sum_{k=0}^{k=T-1} \prod_{i=0}^{i=k} \mathbf{Q}_i)(1)]$$

where T is the period time, $\mathbf{I}$ is identity matrix.

Proof: [15] has given that the expected hitting time in a NHMC is

$$\mathbb{E}(T) = \alpha_B \left[\mathbf{I} + \sum_{n \geq 0}^{\infty} \prod_{k=0}^{n} \mathbf{Q}_t \right] \mathbf{1}_{|B|}, \tag{2}$$

where $|B|$ is the cardinality of set B, $1_{|B|}$ is $|B|$-dimensional column vector whose elements are all 1, α_B is the initial states probability distribution of the network. Note that at the starting time, only source node s owns the message, so the initial state of the network is $Y_0 = 2^{s-1}$, which is the minimum element in set B. Then

$$\alpha_B = \begin{cases} 1 & \text{if } i = 2^{s-1} \\ 0 & \text{if } i \neq 2^{s-1} \end{cases},$$

where $i \in E$. Consequently, the expected hitting time in Chain Y satisfies

$$\begin{aligned} \mathbb{E}(T) &= \left[\mathbf{1}_{|B|} + (\sum_{n \geq 0}^{\infty} \prod_{k=0}^{n} \mathbf{Q}_k) \mathbf{1}_{|B|} \right] (1) \\ &= 1 + (\sum_{n \geq 0}^{\infty} \prod_{k=0}^{n} \mathbf{Q}_t)(1) \end{aligned}, \tag{3}$$

According to network periodicity $\mathbf{Q}_t = \mathbf{Q}_{t+T}$, we can infer that

$$\begin{aligned} \sum_{n \geq 0}^{\infty} \prod_{k=0}^{n} \mathbf{Q}_k &= \mathbf{Q}_0 + \mathbf{Q}_0\mathbf{Q}_1 + \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 + \cdots \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1} \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1}\mathbf{Q}_T + \cdots \\ &= (\mathbf{Q}_0 + \mathbf{Q}_0\mathbf{Q}_1 + \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 + \cdots \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1}) \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1}\mathbf{Q}_0 + \cdots \\ &= (\mathbf{Q}_0 + \mathbf{Q}_0\mathbf{Q}_1 + \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 + \cdots \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1}) \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1}(\mathbf{Q}_0 + \mathbf{Q}_0\mathbf{Q}_1 + \cdots \\ &+ \mathbf{Q}_0\mathbf{Q}_1\mathbf{Q}_2 \cdots \mathbf{Q}_{T-1}) + \cdots \\ &= \sum_{n \geq 0}^{\infty} (\mathbf{Q}_0\mathbf{Q}_1 \cdots \mathbf{Q}_{T-1})^n \sum_{k=0}^{k=T-1} \prod_{i=0}^{i=k} \mathbf{Q}_i. \end{aligned}$$

This gives a significant simplification on expression (3). For the sake of clarity, denote $\mathbf{Q}_0\mathbf{Q}_1 \cdots \mathbf{Q}_{T-1}$ as $\mathbf{R}$. Provided that $(\mathbf{I} - \mathbf{R})$ is nonsingular, $\sum_{n \geq 0}^{\infty} (\mathbf{Q}_0\mathbf{Q}_1 \cdots \mathbf{Q}_{T-1})^n = \sum_{n \geq 0}^{\infty} \mathbf{R}^n = (\mathbf{I} - \mathbf{R})^{-1}$. Then the equation above could be simplified as follows:

$$\sum_{n \geq 0}^{\infty} \prod_{k=0}^{n} \mathbf{Q}_k = (\mathbf{I} - \mathbf{R})^{-1} (\sum_{k=0}^{k=T-1} \prod_{i=0}^{i=k} \mathbf{Q}_i).$$

To this end, we need to prove that $(\mathbf{I} - \mathbf{R})$ is nonsingular matrix. Notice that $\mathbf{Q}_t$ is upper triangular matrix due to $q_{i,j}^t = p_{i,j}^t = 0$ for $j < i$, which provides $\mathbf{R}$ is also upper triangular matrix. As a consequence, $\mathbf{R}$ is nonsingular $\iff$ its determinant $det(\mathbf{I} - \mathbf{R}) \neq 0$. Now that $det(\mathbf{I} - \mathbf{R}) = \prod_{i=2^{s-1}}^{i \in A} (1 - \prod_{k=0}^{T-1} q_{i,i}^k)$ and states $i \in B$(Transient states), $q_{i,i}^k < 1$, $\prod_{k=0}^{T-1} q_{i,i}^k < 1$ and $det(\mathbf{I} - \prod_{i=0}^{i=T-1} \mathbf{Q}_i) > 0$. Hence, $\mathbf{R}$ is nonsingular matrix, which establishes that $(\mathbf{I} - \mathbf{R})$ is nonsingular matrix. ■

C. Homogenous Situation

We progress to discuss the homogenous situation detailedly in this section, that supposes nodes move identically and independently. Throughout this part, the node movement is assumed to be mutually independent and follows a Poisson process, which is validated by many datasets and adopted in many scenarios [2][16]. As a consequence of the definition of homogenous Poisson process, the inter-contact time between two nodes is exponentially distributed with parameter λ. And during each unit interval Δt, the meeting probability of two nodes is $1-e^{-\lambda}$. Because node movement is homogenous, the state transition probabilities of chain Y are stationary. Equivalently stated, chain Y becomes time-homogeneous and $\mathbf{Q}_t$ is equal for all t.

Theorem 2: Under homogenous assumption and epidemic protocol, we have

$$\mathbb{E}(T) \approx \frac{1-e^{-N\lambda}}{\lambda N(e^{\lambda}-1)} + 1 + \frac{1}{2}e^{-\lambda} + \mathrm{O}\left(\lambda e^{-\lambda} N\right). \qquad (4)$$

Proof: Expanding equation (3), we obtained

$$\begin{aligned}\mathbb{E}(T) &= \left[\mathbf{1}_{|B|} + (\sum_{n\geq 0}^{\infty}\prod_{k=0}^{n}\mathbf{Q}_k)\mathbf{1}_{|B|}\right](1)\\ &= 1 + \left[\mathbf{Q}+\mathbf{Q}^2+\mathbf{Q}^3+\cdots\right](1)\\ &= 1 + \mathbf{Q}(1) + \mathbf{Q}^2(1)\cdots\end{aligned}$$

$\mathbf{Q}^n(1)$ represents the sum of the first row elements of matrix $\mathbf{Q}^n$, which is the n-th power of $\mathbf{Q}$. The $\mathbf{Q}$ is shown as (5). Notice that in (5), h means there are totally h nodes infected in the network, including the source node.

For instance, when the network size $N=4$, then $\mathbf{Q}$ would be

$$\begin{bmatrix} e^{-3\lambda} & e^{-2\lambda}(1-e^{-\lambda}) & e^{-2\lambda}(1-e^{-\lambda}) & e^{-\lambda}(1-e^{-\lambda})^2 \\ 0 & e^{-4\lambda} & 0 & e^{-2\lambda}(1-e^{-2\lambda}) \\ 0 & 0 & e^{-4\lambda} & e^{-2\lambda}(1-e^{-2\lambda}) \\ 0 & 0 & 0 & e^{-3\lambda} \end{bmatrix}.$$

The matrix $\mathbf{Q}$ seems to be very complicated, but with proper method, the result of $\mathbb{E}(T)$ could be simplified. First, because $\mathbf{Q}(1)$ equals to the probability that destination node is not infected by the source node - the only node who has the message in the network at the beginning, we have $\mathbf{Q}(1)=e^{-\lambda}$. And,

$$\begin{aligned}\mathbf{Q}^2(1) = &\sum_{i=1}^{i\leq|A|}\mathbf{Q}(1,i)\mathbf{Q}(i)\\ &= e^{-(N-1)\lambda}e^{-\lambda} + \binom{N-2}{1}e^{-(N-2)\lambda}e^{-2\lambda}(1-e^{-\lambda})\\ &+ \binom{N-2}{2}e^{-(N-3)\lambda}e^{-3\lambda}(1-e^{-\lambda})^2+\cdots\\ &+ \binom{N-2}{N-2}e^{-[N-(N-1)]\lambda}e^{-(N-1)\lambda}(1-e^{-\lambda})^{N-2}\\ &= e^{-N\lambda}\left[1+\left(1-e^{-\lambda}\right)\right]^{N-2}.\end{aligned}$$

In this way, we could infer that for $i\geq 1$,

$$\mathbf{Q}^i(1) = e^{-(N+i-2)\lambda}[1+(1-e^{-\lambda})e^{-(i-2)\lambda}]^{N-2}.$$

Therefore,

$$\begin{aligned}\mathbb{E}(T) &= 1+e^{-\lambda}+e^{-N\lambda}\left[1+\left(1-e^{-\lambda}\right)\right]^{N-2}+\cdots\\ &= 1+\sum_{i=0}^{\infty}e^{-(N-1+i)\lambda}\left[1+(1-e^{-\lambda})e^{-(i-1)\lambda}\right]^{N-2}.\end{aligned} \qquad (6)$$

Applying second-order EulerMacLaurin formula with error term [17, 2,p.469], which is

$$\begin{aligned}\sum_{i=0}^{n-1}f(i) = &\int_0^n f(x)\,dx + \frac{B_1}{1!}\left(f(n)-f(0)\right)\\ &+\frac{B_2}{2!}\left(f'(n)-f'(0)\right)\\ &+(-1)^3\frac{1}{2!}\int_0^n B_2(\{x\})f''(x)\,dx\end{aligned}$$

(where B_i is known as the i-th Bernoulli number, $\{x\}=x-\lfloor x\rfloor$, $B_2(\{x\})\in[-1/12,1/6]$), to the equation (6), we get

$$\begin{aligned}\mathbb{E}(T) = 1 &+ \frac{1-e^{-(N-1)\lambda}}{\lambda(N-1)(e^{\lambda}-1)}+\frac{1}{2}e^{-\lambda}\\ &- B_2(\{x\})\lambda e^{-\lambda}[1+(N-2)\left(1-e^{-\lambda}\right)]\\ &\approx \frac{1-e^{-N\lambda}}{\lambda N(e^{\lambda}-1)}+1+\frac{1}{2}e^{-\lambda}\\ &- B_2(\{x\})\lambda e^{-\lambda}[1+N\left(1-e^{-\lambda}\right)]\\ &\approx \frac{1-e^{-N\lambda}}{\lambda N(e^{\lambda}-1)}+1+\frac{1}{2}e^{-\lambda}+\mathrm{O}\left(\lambda e^{-\lambda}N\right).\end{aligned}$$

■

IV. Validation

After introducing our stochastic model, two standard mobility models and one real-life contact trace are used in simulation to assess the accuracy and scalability of our stochastic model. Here we simply compare approximate expected message delay under homogenous assumption to the simulation results, which gives the sufficient validation of our proposed framework. In the following, the datasets are demonstrated, simulation settings and precise results are presented. At last, open problems are discussed as well.

A. Datasets

We perform the simulations on the following three datasets: two traces are generated by simulator based on random waypoint and random direction mobility model, one real-life trace collected in a social network experiment.

1) Random waypoint mobility model: The random waypoint mobility model [18] is one of the most commonly used mobility models in the mobile ad hoc networking research. In the model, Each node of the network chooses an initial location and a destination position uniformly and randomly in a square area, and selects a speed v which is an independent and uniformly random variable within the interval (v_{min}, v_{max}).

$$\begin{bmatrix} e^{-(N-1)\lambda} & e^{-(N-2)\lambda}(1-e^{-\lambda}) & \cdots & e^{-(N-3)\lambda}(1-e^{-\lambda})^2 & \cdots & \cdots & e^{-[N-(N-1)]\lambda}(1-e^{-\lambda})^{N-2} \\ 0 & e^{-2(N-2)\lambda} & \cdots & e^{-2(N-3)\lambda}(1-e^{-2\lambda}) & \cdots & \cdots & e^{-2[N-(N-1)]\lambda}(1-e^{-2\lambda})^{N-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & e^{-(N-1)[N-(N-1)]\lambda}(1-e^{-(N-1)\lambda})^{N-N} \end{bmatrix} \tag{5}$$

The node travels at speed v to the destination and when it reaches the destination, it stays motionless for a predefined duration of time and then selects a new destination point, a new speed value and moves again.

2) Random direction mobility model: In the random direction model, nodes' speed $v \in (v_{min}, v_{max})$ and all nodes move in a square area, which are analogous to the random waypoint model. At the beginning, the nodes choose a speed v, a movement direction $\phi \in (0, 2\pi)$ and certain period τ. Then it keeps moving in the same direction until the predefined time is expired. After that, node repeats the whole process - choosing the new value of random variables and starting movement again. Once travels to the edge of area, it is either reflected or the boundaries wraps around and consequently, the node reappears on the other side [2].

3) UNICAL14 data: The dataset was collected in an experiment at the campus of University of Calabria. There were altogether 15 students participated in the experiment. The experiment lasted one week, from January 28, 2014 to February 5, 2014. During the 7 days, the 15 students' devices detected each other through Bluetooth. The experiment lasted one week, from January 28, 2014 to February 5, 2014, including only the working days. During the experiment, each student's mobile phone was equipped with an Android application called SocialBlueCon, which was able to activate Bluetooth scanning in order to find neighbour devices within 10 meters every 180 seconds [14]. The Bluetooth logs were uploaded to a central server and the dataset was named as UNICAL14. The dataset is small, but up to date, and we also want to test if communication paradigm in opportunistic network can be applied in our campus, which is closer to our daily lives.

B. Simulation Methods and Settings

Simulations are carried out using Opportunistic Networking Environment (ONE) simulator [19]. We simulated the information propagation process under Epidemic routing protocols implemented in ONE[20]. The following settings are used when applying ideal mobility models. Random waypoint and random direction mobility model are generated by ONE, and key parameters are displayed in Table I.

In [2], authors proved that in random waypoint and random direction mobility,

$$\lambda \propto \frac{r\mathbb{E}[V^*]}{L^2},$$

where $\mathbb{E}[V^*]$ is the average speed of nodes and only depends on the predefined parameter (v_{min}, v_{max}). L^2 is the size of movement area. When conducting the experiments using ideal mobility models, we control the velocity and map size in order to make λ only vary with transmit range r.

TABLE I
KEY PARAMETERS

Acronym	Values
World Size(m)	4000, 4000
Walking Speeds(m/s)	1.11, 2.78
Number of Nodes	10 - 70
Upate Interval(s)	1
Transmission Speed(Mbps)	2
Buffer Size(MB)	100
Message Time to Live(h)	20

C. Validation of Assumptions

1) Inter-contact time distribution: The random waypoint and random direction mobility models are generated by the simulator ONE with setting parameters as Table I. The setting of transmit range r is adopted as same as in [2] ($r = 50, 100, 250m$).

Since UNICAL14 data clearly shows 7-days repeating contact pattern (typical weekly pattern), we partitioned the dataset into 7 segments for multiple experiments for analysing students' daily contact pattern more detailedly. Fig. 2 exhibits the distribution of inter-contact time in the project's first day, and figures of inter-contact time distribution in the other six days are omitted here since they are almost same as the Fig. 2.

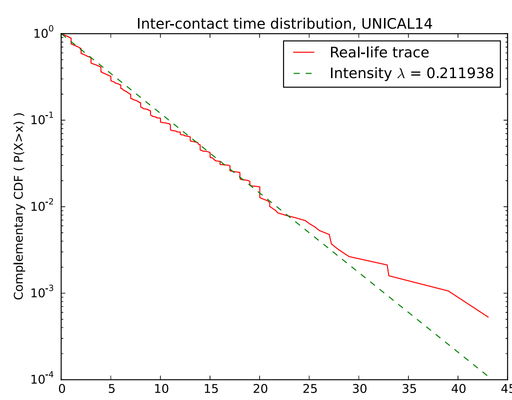

Fig. 2. Inter-contact time distribution for UNICAL14 data

2) Independence of inter-meeting times: In [2], authors already proved that the nodes movements are mutually independent in random waypoint and random direction mobility. So here we merely validate the homogeneous assumption for UNICAL14 dataset. Use the same estimator in[2], which is a classical estimator for the autocorrelation function of $\{\tau(n)_n\}$. And $\{\tau(n)_n\}$ is the n-th inter-contact time between two selected nodes.

$$\rho_n(h) = \frac{\gamma_n(h)}{\gamma_0(h)} \tag{7}$$

where

$$\gamma_n(h) := \frac{1}{n-h}\sum_{i=1}^{n-h}(\tau(i+h) - \hat{\tau}^{(n)})(\tau(i) - \hat{\tau}^{(n)}).$$

We choose a pair of nodes arbitrarily, and the autocorrelation function is plotted in Fig. 3. The autocovariance function is equal to zero for all $h > 1$ if nodes move independently (here, index starts from 1, not 0). Fig. 3 demonstrates that $\rho_n(h)$ that we chose between two stochastic nodes is less than 0.15 and close to zero for the UNICAL14 dataset.

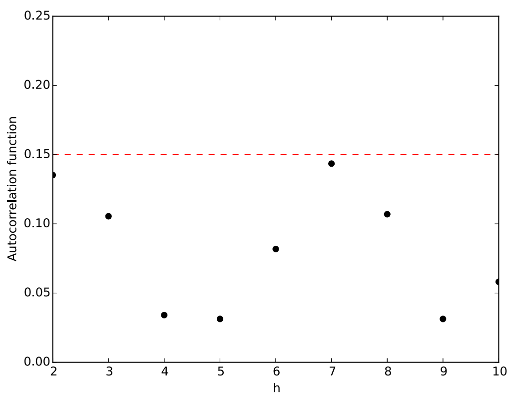

Fig. 3. Autocorrelation function of inter-meeting times for the UNICAL14

D. Results

Fig. 4 presents the simulation and theoretical estimation results based on UNICAL14 dataset. Clearly, the higher intensity λ of mobility process is, the lower the average message latency is. And at the fifth day, students' contacts are the most frequent in one week and λ is the highest. As a consequence, the message delay in fifth day is the lowest, which demonstrates that average message delay $E(T)$ decreases with λ increasing. Our theoretical results are close to the simulation results for the scenario, but not accurate. The inaccuracy is caused by many factors, such as the homogenous assumption does not entirely match real-life situation, and our model lacks consideration of many other factors which will be discussed at the end.

Fig. 5 and Fig. 6 plot the average message delays, including the expected delay and the simulation results, as a function of number of nodes in the RD and RW mobility model. The results shown in the Fig. 5 and Fig. 6 prove that the relationships between average message delay and parameters in our theoretical model match the simulation results well. Then we can conclude from three figures in this section, our model has solid ability to predict the expected message delay in different network scenarios.

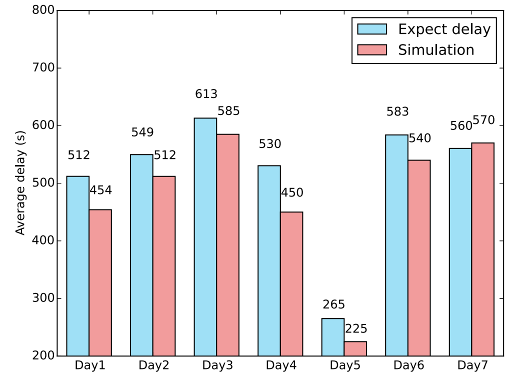

Fig. 4. Message delay in real-life situation

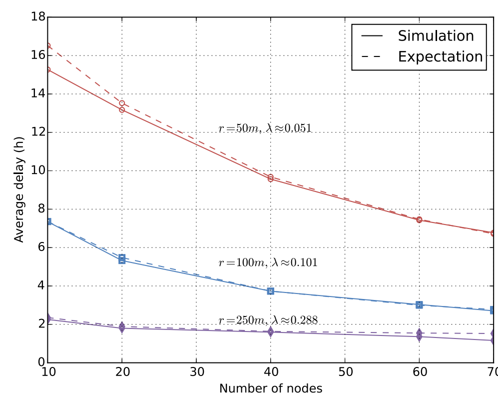

Fig. 5. Message delay in Random Direction Mobility Model

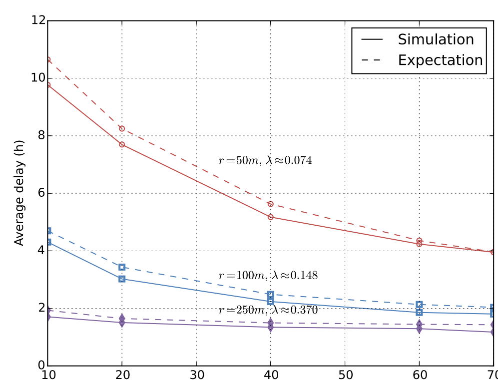

Fig. 6. Message delay in Random Waypoint Mobility Model

E. Discussion

In this part we will briefly discuss about some factors which are not parameters of our analytical framework, but based on empirical observations, they have impact on the message delay and we may need to consider these factors in future work. The results also need to be validated on larger datasets.

In real scenarios, many factors affect the latency gravely, such as social selfishness, communities, message size and so on. For example, the messages in simulations are created uniformly among nodes. But empirically, intra-group nodes have much more frequent contacts than inter-group nodes. Quantifying all these inherent factors is extremely complex and uncompleted. Our model is lack of consideration of these important factors and our assumptions about network are still far from realistic.

Besides, we assume that in each time slot, nodes can only send message to the nodes within the transmission range, that is messages are delivered within one hop. But in real-life situations and system simulations, nodes could send message to nodes that out of regular transmission range through relay nodes in one time slot. Obviously, multi-hop transmission will reduce the message delay gravely. This scheme isn't discussed specifically in this paper, but needs to be further studied.

V. Conclusions & Future work

In this paper, we propose an analytical model based on a 2^{N-1}-state Markov chain for evaluating epidemic routing performance in opportunistic networks. The model focus on link dynamics, which is the main character of message propagation process in Oppnets, and can be applied to different scenarios. Closed-form expressions are derived both under non-homogenous and homogenous assumptions respectively. Particularly, an asymptotic formula is obtained under homogenous situations and can give precise estimation of latency with low computational complexity. At last, simulations are carried out on two standard mobility models, as well as a small real-life trace. Simulation results are compared to our analytical results. All the evidence demonstrates the model's prediction ability and scalability. In the future work, we need to consider more complex routing schemes and network scenarios to build a more robust and comprehensive analytical model.

References

[1] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic networking: data forwarding in disconnected mobile ad hoc networks," *Communications Magazine, IEEE*, vol. 44, no. 11, pp. 134–141, 2006.

[2] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Performance Evaluation*, vol. 62, no. 1, pp. 210–228, 2005.

[3] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance modeling of epidemic routing," *Computer Networks*, vol. 51, no. 10, pp. 2867–2891, 2007.

[4] P. Jacquet, B. Mans, and G. Rodolakis, "Information propagation speed in mobile and delay tolerant networks," *Information Theory, IEEE Transactions on*, vol. 56, no. 10, pp. 5001–5015, 2010.

[5] Y. Li, L. Qiu, D. Jin, L. Su, P. Hui, and L. Zeng, "Contact duration aware evaluation for content dissemination delay in mobile social network," *Wireless Communications and Mobile Computing*, vol. 15, no. 3, pp. 527–537, 2015.

[6] J. Whitbeck, V. Conan, and M. D. de Amorim, "Performance of opportunistic epidemic routing on edge-markovian dynamic graphs," *arXiv preprint arXiv:0909.2119*, 2009.

[7] P. Sermpezis and T. Spyropoulos, "Understanding the effects of social selfishness on the performance of heterogeneous opportunistic networks," *Computer Communications*, vol. 48, pp. 71–83, 2014.

[8] L. Becchetti, A. Clementi, F. Pasquale, G. Resta, P. Santi, and R. Silvestri, "Flooding time in opportunistic networks under power law and exponential intercontact times," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 25, no. 9, pp. 2297–2306, 2014.

[9] C. Boldrini, M. Conti, and A. Passarella, "Performance modelling of opportunistic forwarding under heterogenous mobility," *Computer Communications*, vol. 48, pp. 56–70, 2014.

[10] U. G. Acer, P. Drineas, and A. A. Abouzeid, "Connectivity in time-graphs," *Pervasive and Mobile Computing*, vol. 7, no. 2, pp. 160–171, 2011.

[11] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *Networking, IEEE/ACM Transactions on*, vol. 17, no. 1, pp. 1–14, 2009.

[12] E. Valdano, L. Ferreri, C. Poletto, and V. Colizza, "Analytical computation of the epidemic threshold on temporal networks," *arXiv preprint arXiv:1406.4815*, 2014.

[13] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, "Time-varying graphs and dynamic networks," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 27, no. 5, pp. 387–408, 2012.

[14] A. Socievole, F. De Rango, and A. Caputo, "Wireless contacts, facebook friendships and interests: Analysis of a multi-layer social network in an academic environment," in *Wireless Days (WD), 2014 IFIP*. IEEE, 2014, pp. 1–7.

[15] A. Platis, N. Limnios, and M. Le Du, "Hitting time in a finite non-homogeneous markov chain with applications," *Applied stochastic models and data analysis*, vol. 14, no. 3, pp. 241–253, 1998.

[16] E. Hyytia and J. Ott, "Criticality of large delay tolerant networks via directed continuum percolation in space-time," in *INFOCOM, 2013 Proceedings IEEE*. IEEE, 2013, pp. 320–324.

[17] R. L. Graham, *Concrete mathematics:[a foundation for computer science; dedicated to Leonhard Euler (1707-1783)]*. Pearson Education India, 1994.

[18] C. Bettstetter, G. Resta, and P. Santi, "The node distribution of the random waypoint mobility model for wireless ad hoc networks," *Mobile Computing, IEEE Transactions on*, vol. 2, no. 3, pp. 257–269, 2003.

[19] A. Keränen, J. Ott, and T. Kärkkäinen, "The one simulator for dtn protocol evaluation," in *Proceedings of the 2nd international conference on simulation tools and techniques*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 55.

[20] A. Vahdat, D. Becker *et al.*, "Epidemic routing for partially connected ad hoc networks," Technical Report CS-200006, Duke University, Tech. Rep., 2000.