Modeling Ad Hoc Mobile Networks:

the General k-Hop Relay Routing

Modeling Ad Hoc Mobile Networks: the General k-Hop Relay Routing

Jiajia Liu*‡, Hiroki Nishiyama*, Nei Kato*, Tomoaki Kumagai†, and Atsushi Takahara†
*Tohoku University, Sendai, Japan
†NTT Network Innovation Laboratories, NTT Corporation, Yokosuka, Japan
‡Email: liu-jia@it.ecei.tohoku.ac.jp

***Abstract*—In the last decade, there has been a tremendous increase in both the number of mobile devices and the consumer demand for mobile data communication. As a general network architecture, ad hoc mobile networks are expected to offload a large amount of mobile traffic in lots of promising application scenarios. However, how to achieve a good balance between delivery performances (like delivery delay and delivery probability) and network resource consumptions (like power energy and buffer storage) remains an extremely challenging problem. In this paper, we focus on the general k-hop relay routing, which covers a lot of popular routing schemes as special cases, such as the direct transmission ($k = 1$), the two-hop relay algorithm ($k = 2$), and the epidemic routing ($k = n - 1$). We first develop absorbing continuous-time Markov chain models to characterize the complicated message delivery process under the general k-hop relay routing, and then conduct Markovian analysis to derive all the above important performance metrics. Finally, extensive numerical results are presented to illustrate the achievable delivery performances under the general k-hop relay and the possible performance trade-offs there.**

I. Introduction

In the last decade, there have been a tremendous increase in the number of mobile devices and also a sharp rise in consumer demand for mobile data communication. According to [1], the number of mobile-connected devices would exceed the world's population in 2012, and there will be over 10 billion mobile-connected devices (including machine-to-machine modules) in 2016, approximately 1.4 mobile devices per capita. Furthermore, the mobile data traffic will grow at a compound annual growth rate of 78 percent and increase 18-fold from 2011 to 2016, and by 2016 the monthly global mobile data traffic will surpass 10 exabytes.

As a general network architecture, ad hoc mobile networks are expected to offload a large amount of mobile traffic in lots of promising application scenarios, such as disaster relief, military troop communication, daily information exchange, etc. Note that in these applications, nodes are usually sparsely distributed and it is rather difficult to find an end-to-end routing path at any time instant. For such network scenarios, how to deliver a message from end to end while simultaneously achieving a good balance between delivery performances (like delivery delay and delivery probability) and network resource consumptions (like power energy and buffer storage) remains an extremely challenging problem [2]–[4]. Specifically, adopting too tight a control policy (like a limited message lifetime or restricted relay transmissions) may lead to poor message delivery performances; on the other hand, a loose control policy may result in unnecessary wastes of precious network resources, under which a message may continue to be replicated among relay nodes even after its arrival at the destination.

To address this challenging problem and thus enable a nice trade-off between delivery performances and resource consumptions to be efficiently achieved, a natural choice is to restrict the number of hops that a message (copy) can travel in the network. In this paper, we consider the general k-hop relay routing where each message travels at most k hops to reach the destination after leaving its source. Actually, the k-hop relay routing covers a lot of popular routing schemes as special cases, such as the direct transmission [5] ($k = 1$), the two-hop relay algorithm [6] ($k = 2$), and the epidemic routing [7] ($k = n - 1$), where n is the number of network nodes.

We introduce system models and the general k-hop relay routing in Section II. In Section III, we first develop absorbing continuous-time Markov chain models to characterize the complicated message delivery process under the typical settings of $k = \{1, 2, 3, n - 1\}$, and then conduct Markovian analysis to derive the important delivery performance metrics including the expected delivery delay, the expected delivery cost, and the message delivery probability under any given message lifetime. Different from [8], we provide in this paper general Markovian derivations for delivery delay, cost and probability without deriving embedded discrete-time markov chains. Extensive numerical results are presented in Section IV to illustrate the achievable delivery performances under the general k-hop relay and the possible performance trade-offs there. Finally, we conclude this paper and discuss some future works in Section V.

II. Preliminaries

A. System Models

In this paper, we consider an ad hoc mobile network which consists of n mobile nodes. Without loss of generality, we focus on one source-destination pair, and denote by S and D the source node and the destination node, respectively. Suppose the node S has a single message, say M, to deliver to the node D, and the other $n - 2$ relay nodes have no local traffic to deliver and will act altruistically as pure relays in the delivery process of M. We assume that M can be successfully transmitted during the contact (or meeting) of each node pair.

For the more general cases where M is of message size that cannot be transmitted during a single contact or S has multiple messages to deliver to D, please refer to [9] for details.

We assume that the n mobile nodes move independently in a closed square region, and the time elapsed between two consecutive contacts of any node pair follows an exponential distribution with mean $1/\lambda$, i.e., the occurrence of contacts between any two nodes follows Poisson distribution. The parameter λ is called as inter-meeting intensity, which is related to the network area, node moving speed, and node transmission range [10]. This assumption has been demonstrated to hold nicely for lots of popular mobility models, like Random Direction model and Random Waypoint model, and it has also been widely adopted in literature [11].

B. The General k-Hop Relay Routing

We consider the general k-hop relay routing in this paper, $1 \leq k \leq n-1$, which covers a lot of popular routing schemes as special cases, such as the direct transmission [5] ($k = 1$), the two-hop relay algorithm [6] ($k = 2$), and the epidemic routing [7] ($k = n-1$). Under the k-hop relay routing, each message (or copy) travels at most k hops to reach the destination node after leaving the source node.

When operating under the k-hop relay, each copy of message M (including the original one at the source S and the redundant copies carried by other relay nodes) is labeled with a hop count, which records the total number of hops the message copy has traveled after leaving S. Specifically, the hop count of the original M at S is always 0; when S (resp. a relay carrying M) replicates a new copy of M to other nodes, the new copy is labeled with a hop count which is one hop bigger than that at S (resp. the relay). In other words, for each message replication, if the transmitter carries a copy of M which has a hop count m, $0 \leq m \leq k-1$, the new copy received by the receiver will have a hop count $m+1$. Note that as the maximum hop count is limited to k, only the relay nodes with hop counts less than $k-1$ can further replicate out new copies for M, while those with hop count $k-1$ can only forward the message to the destination D.

There are four basic transmission modes in the k-hop relay:

- "Source-to-Destination" transmission: the source S directly transmits the message M to the destination D.
- "Source-to-Relay" transmission: S meets a relay node and replicates out a new copy of M to the relay node.
- "Relay-to-Relay" transmission: two relay nodes come into contact, and the relay node which carries a copy of M replicates out a new copy of M to the other one.
- "Relay-to-Destination" transmission: a relay node meets the node D and forwards M to D.

Note that for the case $k = 1$, only "Source-to-Destination" transmission is available; for the case $k = 2$, there exist only three basic transmissions, i.e., "Source-to-Destination", "Source-to-Relay", and "Relay-to-Destination".

C. Performance Metrics

For the general k-hop relay routing, we are interested with the following popular performance metrics.

Delivery Delay: for a message M generated at S, the delivery delay is defined as the time elapsed between the time instant when S starts to deliver M and the time instant when the destination D receives M.

Delivery Cost: the delivery cost of M is defined as the total number of transmissions it takes M to travel from S to D.

Delivery Probability: given a message lifetime τ which means M will be dropped from the network if it fails to reach D within time τ, the delivery probability is defined as the probability that D receives M before message expiration.

Reach: it is defined as the fraction of nodes (excluding S) that end up carrying M when D receives M.

III. Markov Chain Models and Markovian Derivations

In this section, we first develop absorbing continuous-time Markov chains to model the delivery process of message M under the general k-hop relay routing, and then conduct Markovian analysis to derive the corresponding delivery delay, delivery cost, delivery probability, and reach.

Before the node D receives M, the network may experience multiple transient states. Note that the transient states are of short duration and they actually represent the possible network states until the delivery process of M becomes absorbed. Before proceeding to develop Markov chain models for the k-hop relay, here we first introduce the transitions that may happen among neighboring network states. With the k-hop relay routing, the possible transition cases are as follows:

- SR transition: it corresponds to the "Source-to-Relay" transmission, and accordingly the number of relay nodes carrying M will be increased by one.
- RR transition: it corresponds to the "Relay-to-Relay" transmission.
- SD or RD transition: it means that either the source S or a relay node successfully delivers M to D, i.e., the "Source-to-Destination" or "Relay-to-Destination" transmission. It will shift the Markov chain into an absorbing state and thus ends the delivery process of M.
- Self-Loop transition: it means that neither the destination node nor a relay node receives M, and thus the network will remain in the current state.

A. The Direct Transmission (Case $k = 1$)

The direct transmission, corresponding to the special setting of $k = 1$, is the simplest case of k-hop relay routing. With such setting, no relay nodes will be employed for helping deliver the message M, and the destination D can only receive M directly from the source S. Therefore, the delivery delay of M can be determined as the time it takes S and D to move into contact, and thus the delivery delay under the direct transmission (case $k = 1$) is exponentially distributed with mean $1/\lambda$. Furthermore, as S will not replicate out any copy

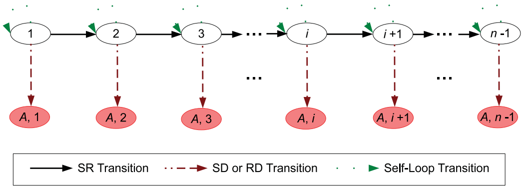
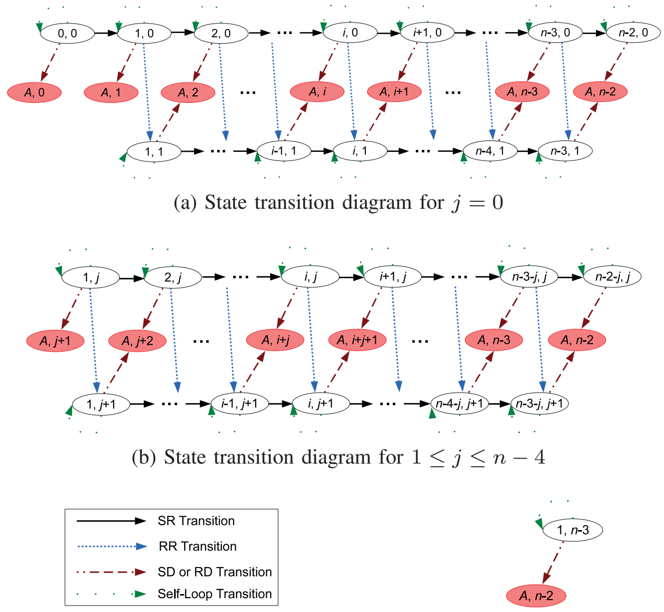


Fig. 1. Absorbing continuous-time Markov chain for two-hop relay routing.



(a) State transition diagram for $j = 0$

(b) State transition diagram for $1 \leq j \leq n-4$

(c) State transition diagram for $j = n-3$

Fig. 2. Absorbing continuous-time Markov chain for three-hop relay routing.

of M to other relay nodes, the corresponding delivery cost is fixed as one.

B. The Two-Hop Relay (Case $k = 2$)

Under the two-hop relay routing (case $k = 2$), D either directly receives M from S or receives M from a relay node which has received a copy of M from S before. A relay node carrying M cannot replicate M to other relay nodes except forwarding M to D. Therefore, no RR transition will happen during the delivery process of M.

We denote by A the absorbing state that M has been received by D, and denote by i a network state that there are i copies of M (including the original one at S) in the network while D has not received M yet, $1 \leq i \leq n-1$. Suppose the network system is within state i, $1 \leq i \leq n-2$, then we can see that it may transit to different states under different transition cases. Specifically, it will transit to state $i+1$ (resp. state i) under the SR transition (resp. the Self-loop transition), and it will enter state A under the SD or RD transition.

Since the network may enter absorbing state A via different transient states, we denote by state (A, i) that the network becomes absorbed via transient state i, $1 \leq i \leq n-1$. After combining the possible transitions of all states i, $1 \leq i \leq n-1$, we obtain an absorbing continuous-time Markov chain as illustrated in Fig. 1. One can easily observe from Fig. 1 that there are in total $n-1$ transient states and $n-1$ absorbing states. Given the Markov chain in Fig. 1 gets absorbed in state (A, i), one can see that there are $i-1$ relay nodes carrying message M when D receives M. Thus, the corresponding delivery cost and reach of message M can be determined as i and i/n, respectively.

C. The Three-Hop Relay (Case $k = 3$)

With the three-hop relay routing, only the source S and the relay nodes receiving M directly from S, can replicate M to other nodes; a relay node which has received a copy of M from another relay rather than from S, can only forward M to the destination D. One can easily see that in the above delivery process, there exist the following two kinds of relay nodes: a relay node is called as a tier 1 relay if it receives M directly from S; otherwise, if it receives M from a tier 1 relay, it is called as a tier 2 relay. Obviously, by restricting the replication of M to S and tier 1 relay nodes, M travels at most three hops to reach D.

Now we proceed to characterize the delivery process of M under the three-hop relay routing. We denote by (i, j) a transient state that there are i tier 1 relay nodes and j tier 2 relay nodes in the network, $0 \leq i \leq n-2$, $0 \leq j \leq n-2$, $0 \leq i+j \leq n-2$, and denote by (A, m) an absorbing state that when D receives M there are m relay nodes carrying M, $0 \leq m \leq n-2$. Suppose the network is within state (i, j), the transitions that may happen are as follows: 1) SR transition where the number of tier 1 relay nodes will be increased by one and the network will transit to state $(i+1, j)$; 2) RR transition which will increase the number of tier 2 relay nodes by one and shift the network into state $(i, j+1)$; 3) SD or RD transition which will finish the delivery process and thus shift the network into absorbing state $(A, i+j)$; 4) Self-Loop transition under which the network will transit from state (i, j) back to itself. After combining the transition cases of all transient states (i, j), the delivery process of M can be defined by a finite-state absorbing Markov chain as shown in Fig. 2.

Note that for the Markov chain in Fig. 2, Figs. 2a, 2b, and 2c each represents some cases of the full Markov chain. Specifically, Fig. 2a defines the transitions among neighboring states where there is no more than one tier 2 relay node, i.e., $j = 0$; Fig. 2b represents the cases that tier 1 relay nodes may deliver M to at most one more relay given that there are already j tier 2 relay nodes, $1 \leq j \leq n-4$; Fig. 2c shows how the node D may receive M when there are $n-3$ tier 2 relay nodes. One can easily observe from Fig. 2 that there are actually $n-2$ rows of transient states. If we denote by L_m the number of transient states in the m_{th} row, $0 \leq m \leq n-3$, then we have $L_m = n-1$ when $m = 0$ and $L_m = n-2-m$ when $1 \leq m \leq n-3$. Therefore, the total number of transient states in the Markov chain of Fig. 2 can be calculated as $\frac{1}{2}(n^2 - 3n + 4)$. It is further noticed that there are in total $n-1$ absorbing states (A, m), $0 \leq m \leq n-2$. Given the Markov chain of Fig. 2 becomes absorbed in state (A, m), the corresponding delivery cost of M can be easily determined as

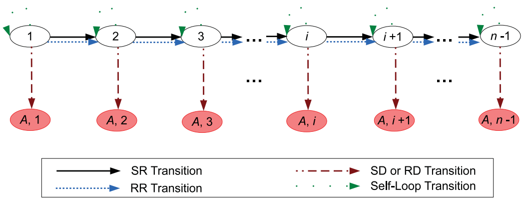


Fig. 3. Absorbing continuous-time Markov chain for the epidemic routing.

$m+1$, $0 \leq m \leq n-2$.

D. The Epidemic Routing (Case $k = n-1$)

For the considered network which consists of one source-destination pair and $n-2$ relay nodes, it takes M at most $n-1$ hops to reach D after leaving S. Therefore, there is actually no limitation on the message replication behavior with $k = n-1$, i.e., any node carrying M, no matter S or relay nodes, can replicate out new copies of M to other relay nodes. A relay node newly receiving M can further replicate it to other relay nodes, resulting a message dissemination process identical to the spreading of infectious diseases. Thus, the general k-hop relay with $k = n-1$ is also known as epidemic routing.

If we use state i to denote a transient state that there are i copies of M in the network, $1 \leq i \leq n-1$, and denote by state (A, i) the absorbing state that when D receives M there are i mobile nodes carrying M, i.e., the last transient state before absorption is state i, the delivery process of M under the epidemic routing can then be defined by an absorbing Markov chain shown in Fig. 3. Note that different from the case $k = 2$ in Fig. 1, there exist RR transitions between neighboring transient states i and $i+1$ in Fig. 3, $1 \leq i \leq n-2$.

E. Markovian Derivations

For message M, we use T_d, C_d, and $P_d(\tau)$ to denote the delivery delay, the delivery cost, and the delivery probability under message lifetime τ, respectively. Based on the Markov chain models developed in Section III, we now proceed to conduct Markovian analysis to derive $\mathbb{E}\{T_d\}$, $\mathbb{E}\{C_d\}$, and $P_d(\tau)$ for the general k-hop relay routing.

Suppose in the continuous-time Markov chain (CTMC), there are in total β transient states and α absorbing states. We number the β transient states sequentially as $1, 2, \ldots, \beta$ in a left-to-right and top-to-down way, and number the α absorbing states sequentially as $1, 2, \ldots, \alpha$ similarly. After combining the transition rates among all transient states in the absorbing CTMC, we are able to define a transition rate matrix $\mathbf{Q} = (q_{ij})_{\beta \times \beta}$, where the entry q_{ij} denotes the rate of transiting from transient state i to transient state j, $i, j \in [1, \beta]$. Note that if there is no transition connecting states i and j, the corresponding ij-entry in $\mathbf{Q}$ is 0. For each transient state in the CTMC, the rate of transiting back to itself is always negative. Therefore, all the main diagonal entries in $\mathbf{Q}$, i.e., $\{q_{ii}\}$, $1 \leq i \leq \beta$, are always of negative values.

After defining the transition rate matrix $\mathbf{Q}$, the expected delivery delay of M, i.e., $\mathbb{E}\{T_d\}$, can be determined as

$$\mathbb{E}\{T_d\} = -\mathbf{e} \cdot \mathbf{Q}^{-1} \cdot \mathbf{c}, \tag{1}$$

where $\mathbf{Q}^{-1}$ is the inverse matrix of $\mathbf{Q}$, the ij-entry of $-\mathbf{Q}^{-1}$ denotes the duration of network staying in state j given that it starts from state i, $\mathbf{e}$ is the initial row vector $\mathbf{e} = (1, 0, \ldots, 0)$, and $\mathbf{c}$ is a column vector $\mathbf{c} = (1, 1, \ldots, 1)^T$ [12].

If M is associated with a limited lifetime τ, the delivery probability $P_d(\tau)$ can be given by

$$P_d(\tau) = 1 - \mathbf{e} \cdot e^{\tau \cdot \mathbf{Q}} \cdot \mathbf{c}, \tag{2}$$

where the matrix exponential

$$e^{\tau \cdot \mathbf{Q}} = \sum_{m=0}^{\infty} \frac{\tau^m}{m!} \mathbf{Q}^m. \tag{3}$$

Since there exists a delivery cost, say h_m, when the CTMC enters an absorbing state m, $m \in [1, \alpha]$, we can accordingly define a column vector $\mathbf{h} = (h_m)_{\alpha \times 1}$. One can see that in order to get the expected delivery cost $\mathbb{E}\{C_d\}$, the only remaining issue is to derive the probability of the CTMC entering each absorbing state m, $m \in [1, \alpha]$. If we use v_{lm} to denote the rate of the CTMC transiting from transient state l to absorbing state m, and use v_{ll} to denote the rate of the CTMC transiting from transient state l back to itself, $l \in [1, \beta]$, $m \in [1, \alpha]$, then we have the one-step transition probability matrix $\mathbf{R} = (p_{lm})_{\beta \times \alpha}$ where the lm-entry p_{lm} is given by

$$p_{lm} = -\frac{v_{lm}}{v_{ll}}. \tag{4}$$

With $\mathbf{R}$ denoting the one-step probabilities of the CTMC transiting from the β transient states to the α absorbing states, $\mathbb{E}\{C_d\}$ can be given by

$$\mathbb{E}\{C_d\} = \mathbf{e} \cdot \mathbf{Q}^{-1} \cdot \mathbf{D} \cdot \mathbf{R} \cdot \mathbf{h}, \tag{5}$$

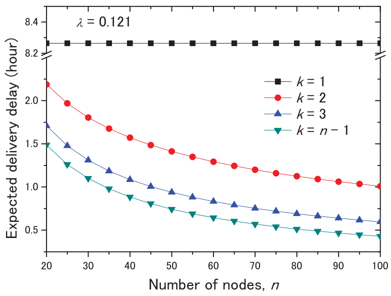
where $\mathbf{D}$ is the diagonal matrix with diagonal entry $d_{ii} = v_{ii}$, and the ij-entry of matrix multiplication $\mathbf{Q}^{-1} \cdot \mathbf{D} \cdot \mathbf{R}$ denotes the probability of the CTMC entering absorbing state j given that it starts from transient state i, $1 \leq i \leq \beta$, $1 \leq j \leq \alpha$.

As there is a new node receiving a copy of M after each transmission, the expected reach can be therefore determined as $\mathbb{E}\{C_d\}/n$.

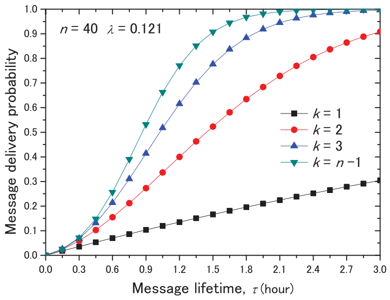
IV. Numerical Results

As illustrative examples, we fix λ as $\lambda = 0.121$ (contact / hour), let n vary from 20 to 100, and summarize in Figs. 4 and 5 the achievable performances of the general k-hop relay routing. Note that the numerical results under other settings of λ can also be easily obtained by our Markovian derivations as well.

One can easily observe from Figs. 4a and 4b that by allowing more hops (i.e., a bigger k) in the message delivery process, both the expected delivery delay and delivery probability can be significantly improved. However, it is also noticed that as k varies from 3 to $n-1$, the performance improvements of both delivery delay and delivery probability are actually rather limited. A further careful observation of Fig. 4a indicates that as n increases, the performance gap between the settings of $k = 3$ and $k = n-1$ remains almost unchanged. Figs. 5a and 5b show how the expected delivery

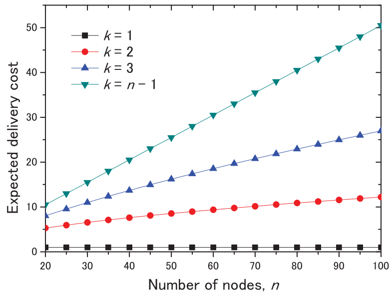


(a) $\mathbb{E}\{T_d\}$ vs. n

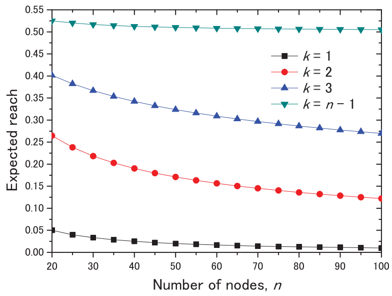


(b) $P_d(\tau)$ vs. τ

Fig. 4. $\mathbb{E}\{T_d\}$ and $P_d(\tau)$ under the settings of $k = \{1, 2, 3, n-1\}$.



(a) $\mathbb{E}\{C_d\}$ vs. n



(b) Expected reach vs. n

Fig. 5. $\mathbb{E}\{C_d\}$ and expected reach under the settings of $k = \{1, 2, 3, n-1\}$.

cost and expected reach vary with n, respectively. We can see that the gap between the curve of $k = n-1$ and that of $k = 3$ in both Figs. 5a and 5b, rises up sharply as n increases from 20 to 100. In light of the limited performance gains of both expected delivery delay and delivery probability observed in Fig. 4, a moderate value of k (e.g., $k = 3$ in Figs. 4 and 5, which may depend on the actual network settings) may be efficient enough to achieve a flexible trade-off between the delivery performances and energy consumptions.

V. Conclusion

In this paper, we have investigated the popular delivery performances for the general k-hop relay routing in ad hoc mobile networks, like the delivery delay, the delivery probability, the delivery cost, and average reach. Absorbing continuous-time Markov chains have been developed to characterize the message delivery process of the k-hop relay routing under the typical settings $k = \{1, 2, 3, n-1\}$. Extensive numerical results have been further presented to show the achievable delivery performances and the possible performance trade-offs there. Our results indicate that the performance gains between the settings of $k = 3$ and $k = n-1$ could be rather limited, and a moderate value of k can be efficient enough to achieve a flexible trade-off between the delivery performances and energy consumptions. One interesting future direction is to extend the Markov chain models developed in this paper to other settings of k, $3 < k < n-1$.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2011-2016," White Paper, February 2012.

[2] A. Krifa, C. Barakat, and T. Spyropoulos, "Message drop and scheduling in dtns: Theory and practice," *IEEE Transactions on Mobile Computing*, vol. 11, no. 9, pp. 1470–1483, September 2012.

[3] L. X. Cai, L. Cai, X. Shen, and J. W. Mark, "Rex: a randomized exclusive region based scheduling scheme for mmwave wpans," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 113–121, January 2010.

[4] P. Thulasiraman, J. Chen, and X. S. Shen, "Multipath routing and max-min fair qos provisioning under interference constraints in wireless multihop networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 5, pp. 716–728, May 2011.

[5] T. Matsuda and T. Takine, "(p, q)-epidemic routing for sparsely populated mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 783–793, June 2008.

[6] M. Grossglauser and D. N. Tse, "Mobility increases the capacity of ad hoc wireless networks," in *INFOCOM*, 2001.

[7] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," in *Duke Univ., Durham, NC, Tech.Rep.CS-200006*, April 2000.

[8] J. Liu, H. Nishiyama, and N. Kato, "Performance modeling of three-hop relay routing in intermittently connected mobile networks," in *WCNC*, 2013.

[9] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, "On the delivery probability of two-hop relay manets with erasure coding," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1314–1326, April 2013.

[10] R. Groenevelt, "Stochastic models in mobile ad hoc networks," Ph.D. dissertation, University of Nice Sophia Antipolis, April 2005.

[11] E. Altman, T. Basar, and F. D. Pellegrini, "Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks," in *Inter-Perf*, 2008.

[12] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. D. Van Nostrand, 1963.