On Exploiting Contact Patterns for Data Forwarding in Duty-Cycle Opportunistic Mobile Networks

Huan Zhou, Jiming Chen, *Senior Member, IEEE*, Hongyang Zhao, Wei Gao, *Member, IEEE*, and Peng Cheng, *Member, IEEE*

***Abstract*—In this paper, we focus on investigating the impact of duty-cycle operation on data forwarding in duty-cycle opportunistic mobile networks (OppNets) and on designing an efficient data-forwarding strategy for duty-cycle OppNets. Some recent studies utilize node contact patterns to aid in the design of a data-forwarding strategy in OppNets. However, when duty-cycle operation is applied in OppNets, several node contacts will be missed when nodes are in the sleep state for energy saving, and it becomes challenging to design an efficient data-forwarding strategy based on exploitation of node contact patterns. To address this challenge, we first propose a model to investigate the contact process in duty-cycle OppNets and to estimate the probability of contact discovery. We also experimentally validate the correctness of our proposed model. Second, based on this model, we propose a novel approach to improve the performance of data forwarding in duty-cycle OppNets. The proposed forwarding strategy takes into account both the contact frequency and contact duration and manages to forward data copies along the opportunistic forwarding paths, which maximize the data delivery probability. Finally, extensive real-trace-driven simulations are conducted to compare the proposed data-forwarding strategy with other recently reported data-forwarding strategies in terms of delivery ratio and cost. The simulation results show that our proposed data-forwarding strategy is close to the Epidemic Routing strategy in terms of delivery ratio but with significantly reduced delivery cost. Additionally, our proposed strategy outperforms the Bubble Rap and Prophet strategies in terms of delivery ratio with reasonable delivery cost.**

***Index Terms*—Data forwarding, duty-cycle operation, node contact pattern, opportunistic mobile networks (OppNets).**

I. Introduction

RECENTLY, with the rapid proliferation of portable devices (e.g., personal digital assistants and smartphones), a new peer-to-peer application scenario, i.e., opportunistic mobile networks (OppNets), has begun to emerge [1]–[7]. In OppNets, it is hard to maintain end-to-end communication links due to the time-varying network topology, and only intermittent connectivity among portable devices (called nodes below) exist in the network. Nodes have to store their data to be transmitted and opportunistically forward data upon contacts with others. This communication paradigm is referred to as the store–carry–forward mechanism, standing as a basic strategy of data delivery in OppNets.

To enable this data delivery, nodes have to be kept in idle listening mode to discover if there is any neighbor node in its vicinity. In OppNets, the intercontact time is generally much longer than the contact duration due to node sparsity; hence, nodes will spend most of their energy in idle listening mode during intercontact times. Experimental studies in [8]–[10] have shown that energy consumption in idle listening mode is almost as much as that in a receiving mode. Therefore, over 95% of the node energy is consumed by the idle listening mode searching for neighbors [11]. This makes energy saving an important problem with OppNets; it consists of nodes with limited power supply.

Duty cycle is an effective approach to energy saving and enables mobile nodes to alternatively operate between wake-up and sleep states. The duty-cycle operation can be classified into two categories: synchronous and asynchronous [12], [13]. Since synchronous duty-cycle operation requires global time synchronization, which leads to unacceptable overhead of global communication, the asynchronous operation is more favorable in OppNets. Duty-cycle operation also significantly reduces the performance of data transmission because nodes miss a large amount of contact opportunities when they switch to the sleep state. Therefore, it is pressing to investigate the impact of duty-cycle operation on the performance of data transmission in OppNets.

In this paper, we focus on investigating the impact of duty-cycle operation on data forwarding in OppNets. Data forwarding has been well studied in OppNets, and several data-forwarding strategies have been proposed. Some data-forwarding strategies rely on comparisons between nodes' probabilities of contacting the destination to make forwarding decisions [14]–[16]. However, the performance of these schemes is limited due to the lack of global information at individual nodes about how to reach the destination. Recently, some studies have exploited the characteristics of node contact patterns, and data-forwarding decisions are decided based on nodes' cumulative contact characteristics over a long period of time [3], [17]–[20]. Since node contact patterns represent the

Manuscript received December 25, 2012; revised April 19, 2013; accepted May 24, 2013. Date of publication June 10, 2013; date of current version November 6, 2013. This work was supported in part by the National Natural Science Foundation of China under Grant 61061130563, Grant 61004060, Grant 61222305, and Grant 61174177; by the 863 High-Tech Project under Grant 2011AA040101-1; by the National Technology Research and Development Program under Grant 2013AA10230207; by the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20100101110066, Grant 20120101110139, and Grant NCET-11-0445; and by the Fundamental Research Funds for the Central Universities under Grant 2013QNA5013 and Grant 2013FZA5007. The review of this paper was coordinated by Prof. G. Mao. *(Corresponding author: P. Cheng.)*

H. Zhou, J. Chen, H. Zhao, and P. Cheng are with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: pcheng@iipc.zju.edu.cn).

W. Gao is with the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2013.2267236

long-term relation among nodes with better stability, they make a data-forwarding decision more effective and less susceptible to the randomness of node mobility.

However, in duty-cycle OppNets, since nodes in the network operate at a duty cycle and switch between the sleep and wake-up states, several contacts will be missed when nodes switch to the sleep state. *This leads to distinct characteristics of the contact process in duty-cycle OppNets* and *makes it difficult to utilize node contact patterns to design efficient data-forwarding strategy.* To solve this problem, we propose a model to investigate the characteristics of the contact process in duty-cycle OppNets. For simplicity, we consider periodic duty cycling of individual nodes, such that each node independently determines the start time of itself and works for a constant time T_{on} every period T. We quantify the contact discovery probability as a function of wake-up time T_{on}, period T, and contact duration T_d. Here, the contact discovery probability represents the probability that a contact between two nodes can be discovered by each other in duty-cycle OppNets. By doing so, we manage to analytically characterize the relationship between energy consumption and the contact discovery probability. Based on the proposed model, we then propose a novel approach to data forwarding in duty-cycle OppNets. The proposed data-forwarding strategy takes into account both the contact frequency and the contact duration, and manages to forward data copies toward the opportunistic forwarding paths, which maximize the data delivery probability. Our contributions in this paper are fourfold.

1) We propose a model to investigate the contact process in duty-cycle OppNets and derive the contact discovery probability. Given that contact duration follows the power law distribution, we analytically explore the relationship between energy consumption and contact discovery probability.
2) We use real mobility traces to validate the correctness of our proposed model. Via real-trace-driven simulations, our results show that the simulation results are quit close to the theoretical results, which validate the correctness of our proposed model.
3) Based on the proposed model, we propose a novel approach to improve the performance of data forwarding in duty-cycle OppNets by exploiting node contact patterns in duty-cycle OppNets.
4) Extensive real-trace-driven simulations are conducted to evaluate the performance of our proposed data-forwarding strategy. The simulation results show that our proposed strategy outperforms other recently reported strategies under the considered scenarios.

The remainder of this paper is organized as follows. We present the related work in Section II and give the network model in Section III. Section IV derives the contact discovery probability and analyzes the relationship between energy consumption and the contact discovery probability. We perform real-trace-driven simulations in Section V to validate the correctness of our proposed model in Section IV. Based on the proposed model in Section IV, we present our proposed data-forwarding strategy in Section VI; then, extensive real-trace-driven simulations are conducted in Section VII to evaluate the performance of our proposed data-forwarding strategy. We conclude this paper in Section VIII.

II. Related Work

Here, we first introduce the related work about energy-saving mechanisms in OppNets and then introduce the related work about data-forwarding strategies in OppNets.

A. Energy-Saving Mechanisms in OppNets

Several energy-saving mechanisms have been proposed in OppNets. Since nodes cost much energy in the contact-probing process and a high probing frequency means a large amount of energy consumption, some studies have investigated the contact-probing process to save energy [21]–[24] in OppNets. The impact of contact probing on the probability of missing a contact and the tradeoff between the missing probability and energy consumption in Bluetooth devices were investigated in [21]. Real trace-driven simulation results show that their proposed adaptive contact-probing mechanism, i.e., STAR, consumes three times less energy when compared with a constant contact-probing interval scheme. The impact of contact probing on link duration and the tradeoff between the energy consumption and throughput were investigated in [22]. In addition, this paper provides a framework for computing the optimal contact-probing frequency under energy limitations and adjusts the probing frequency according to the node-encountering rate. Two novel adaptive schemes for dynamically selecting the parameters of the contact-probing process were introduced and evaluated in [23]. Nodes in the network switches between two radios: a low-power radio, with slow discovery mode for discovering contacts, and a high-power radio, with fast discovery mode for transmitting data, depending on a mobility context. Simulation results show that their adaptive algorithms can reduce energy consumption by 50% and have up to 8% better performance over a static power-conserving scheme.

However, all aforementioned works only focus on investigating the contact-probing process to save energy; they do not take into account the energy consumption in idle listening mode. The idle listening mode consumes much more energy than the contact-probing process. Therefore, some works have applied the duty-cycle operation to save energy in OppNets. Stationary battery-powered nodes, which are called throw-boxes, were used in [11] to enhance the capacity of OppNets. This work presents a duty-cycled controller for long-range radios that predicts when and for how long the mobile node will be in the range of the throw-box, and this paper builds a bus-based delay-tolerant network test bed to test their proposal. However, the proposal needs beacon position, speed, and direction to feed the prediction algorithm, and also needs GPS data to obtain time synchronization, which is difficult to realize for OppNets consisting of portable devices with limited energy supplies. Power saving tradeoffs in duty-cycle OppNets as a function of the wake-up and sleep intervals and as a function of node contact duration were investigated in [25]. This work investigates the tradeoff between energy saving and contact discovery

probability, as well as the tradeoff between delay-tolerant object dissemination time and energy saving in duty-cycle OppNets. However, this work only gives the tradeoff when the contact duration is a certain value. Another work [26] also proposed a model to investigate the tradeoff between energy saving and contact discovery probability in duty-cycle OppNets. This work gives the tradeoff when the contact duration follows the power law distribution and uses real mobility trace to validate their proposed model.

B. Data-Forwarding Strategies in OppNets

In the previous works, several data-forwarding strategies have been proposed in OppNets. Epidemic routing [27] is a widely used data-forwarding strategy in OppNets, which simply floods data to the entire network. This strategy can guarantee a high data delivery ratio but is expensive in terms of delivery cost since data in the network are essentially flooded. Attempts to reduce the delivery cost are explored in [28] and [29]. A simple approach to reduce the delivery cost of flooding by only forwarding a copy of data with some probability p $(p < 1)$ was proposed in [28]. The spray-and-wait scheme proposed in [29] reduces the delivery cost by assigning a small number of replica copies to a data item and distributes data copies to a number of relay nodes by the source node, and then waits until a relay node meets the destination. Some other works introduce a destination-based approach to reduce the delivery cost and to increase the data delivery performance, whereas the metrics are derived from calculating the probability of delivery to the destination node. FRESH [14] uses the time elapsed since the last contact with the destination node as the data-forwarding metric. Prophet [15] calculates the probability by using the past contact histories to predict the probability of meeting a node again, and data copies are forwarded to nodes that have higher contact probability for the destination node. To reduce delivery costs even more, delegation forwarding in [16] seeks to forward data copies only to nodes whose quality metric is the highest so far.

Recently, some works [3] [18]–[20] utilize node contact patterns to design efficient data-forwarding strategies for OppNets. For example, cumulative contact probability (CCP) as the centrality metric for multicasting in OppNets is proposed in [3], whereas CCP is calculated based on the cumulative node contact frequency and the assumption of exponential distribution of pairwise node intercontact time. SimBet [19] uses "betweenness" centrality metrics and social similarity to increase the performance of data forwarding, in which data copies are forwarded toward the node with both higher centrality and similarity to increase the possibility of finding the potential forwarder to the destination node. LABEL [20] uses a small label to indicate affiliation information to help data forwarding in OppNets, in which data copies are forwarded to nodes that are in the same community with the destination node. This work proves the intuition that simply identifying a community can improve data delivery, even during a conference where the people from different subcommunities tend to mix together. Bubble Rap [18] considers both the community in LABEL and the node centrality information in SimBet to increase the data-forwarding performance. Data copies are first forwarded to nodes that have higher global centrality. When data copies are forwarded to the same community as the destination, then local centrality will be used instead of global centrality as the forwarding metrics, and data copies are continued to bubble up until the destination is reached or the data are expired. However, these existing strategies are not suitable to the duty-cycle OppNets; this is because the characteristics of the contact process in duty-cycle OppNets are obviously different from that in OppNets, which makes it difficult to analyze node contact patterns in duty-cycle OppNets.

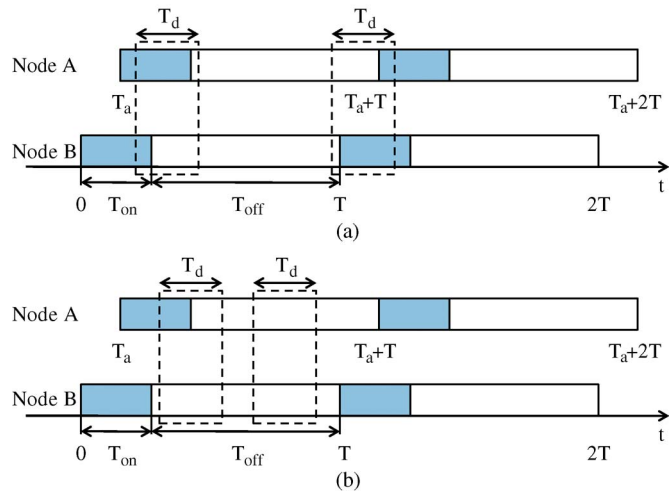


Fig. 1. Contact between two nodes in duty-cycle OppNets. Each node switches from the sleep state to the wake-up state at the beginning of T_{on}. (a) Effective contact. (b) Missed contact.

This paper differs from the previously stated data-forwarding strategies because our approach proposes a model to investigate the characteristics of the contact process in duty-cycle OppNets. Hence, it is more suitable for the duty-cycle OppNets than other data-forwarding strategies designed in OppNets. Furthermore, our proposed data-forwarding strategy utilizes node contact patterns to aid the design of an efficient data-forwarding strategy. Therefore, our proposed data-forwarding strategy is more efficient and suitable for data forwarding in duty-cycle OppNets.

III. Network Model

Without loss of generality, we assume that nodes in the network have enough buffer size to store data and that the size of each data is small enough so that nodes can complete the data exchange process in each contact. Similar to the network model in [25], each node in the network is duty-cycled with two states, i.e., the wake-up state and the sleep state, as shown in Fig. 1.

1) Wake-up state: A node is in the wake-up state during an interval of time equal to T_{on}. Nodes in the wake-up state can exchange data with other nodes, send beacon messages periodically to discover contacts with other nodes, or listen to the wireless channel to discover beacon messages from other nodes.
2) Sleep state: a node is in the sleep state during the remaining interval of the period T that is equal to T_{off}. Nodes in

the sleep state switch their wireless interfaces off to save energy; thus, they cannot communicate with other nodes.

Note that $T = T_{\text{on}} + T_{\text{off}}$; thus, the duty cycle will be (T_{on}/T). Since nodes switch from the sleep state to the wake-up state at the beginning of T_{on}, after choosing a random time to start, nodes in the network will schedule themselves to be in the wake-up state for constant time T_{on} every period T to save energy. A certain node in the wake-up state can periodically send beacon messages to discover contacts with other nodes; all nodes in the communication range that hear the beacon message respond to this node with some information (e.g., identity, services available, etc.). Based on this information, this node can record the contact history with its neighbor nodes.

In OppNets, two nodes contact each other if they are within the communication range of each other, and the interval when nodes are continuously in contact with each other is called the contact duration. Fig. 1 gives an example about the contact between two nodes in duty-cycle OppNets. As shown in Fig. 1, a contact between two nodes A and B happens at the beginning of T_d and lasts for T_d, which represents the contact duration. Let T_B be a random variable indicating the time at which node B switches from the sleep state to the wake-up state at a certain period, and let T_A be a random variable indicating the time at which node A switches from the sleep state to the wake-up state at the same period. To facilitate the modeling, as shown in Fig. 1, we make node B to switch from the sleep state to the wake-up state at time 0 in this period, and we let T_a be a random variable indicating the offset time between nodes A and B, which is equal to $|T_A - T_B|$. Note that, in duty-cycle OppNets, nodes may miss contacts with other nodes when they switch from the wake-up state to the sleep state to save energy. Therefore, we divide the contact in duty-cycle OppNets into two kinds: the *effective contact* and the *missed contact*. The effective contact contains two cases. The first case happens when two nodes are both in the wake-up state at the beginning of their contact. The second case happens when two nodes are not both in the wake-up state at the beginning of their contact, but they will be both in the wake-up state before the contact ends, as shown in Fig. 1(a). Since this kind of contact between two nodes can be discovered by each other during the contact with each other, we thus regard this kind of contact as the effective contact, which can be used for data exchange. The missed contact happens when two nodes are not both in the wake-up state during contact with each other, as shown in Fig. 1(b). Since this kind of contact between two nodes cannot be discovered by one another, we thus refer to this kind of contact as missed contact. Note that the contact in OppNets is infrequent, and the contact process has a significant effect on data forwarding in OppNets. Therefore, in the following, we will propose a model to investigate the contact process in duty-cycle OppNets and to analyze the relationship between energy consumption and contact discovery probability under different situations.

IV. Modeling the Contact Process in Duty-Cycle OppNets

In OppNets, unlike traditional connected networks (e.g., peer-to-peer networks and Internet-accessible networks), nodes are intermittently connected. In the previous studies, authors in [3], [17], and [18] found that the characteristics of contact process in OppNets follow a certain regularity. However, in duty-cycle OppNets, since nodes in the network operate at a duty cycle and switch between the sleep and wake-up states, several contacts will be missed when nodes switch to the sleep state. This leads to distinct characteristics of a contact process in duty-cycle OppNets, and this makes it difficult to utilize node contact patterns to design an efficient data-forwarding strategy. Therefore, here, we propose a model to investigate the contact process in duty-cycle OppNets and analyze the relationship between energy consumption and contact discovery probability.

A. Contact Discovery Probability

Here, we give the definition and expression of the contact discovery probability. Nodes in the network wake up and sleep asynchronously, as introduced in Section III. Nodes A and B switch from the sleep state to the wake-up state at time T_A and T_B at a certain period, respectively. T_a indicates the offset time between nodes A and B, which is equal to $|T_A - T_B|$. Since T_A and T_B are independent of each other and each node in the network works for a constant time T_{on} at every period T, then we can obtain that the offset time T_a is uniformly distributed over the period T, which can be expressed as

$$f_a(t) = \frac{1}{T} \quad \text{for} \quad 0 \le t \le T \tag{1}$$

where $f_a(t)$ is the probability distribution function (pdf) of T_a.

Let T_{ct} be a random variable indicating the time when a contact would begin. As shown in Fig. 1(a), T_{ct} can be expressed as the beginning of T_d. Note that authors in [25] state that "in sparse networks in which the contact rates are low, the intercontact time is higher than the period T." Since the contact rates in OppNets are also low, e.g., the average intercontact time in the *Infocom 06* trace [30] is 125.95 min, whereas in the *MIT Reality,* trace [31] is 334.2 h (or 13.926 days), we also assume that the intercontact time in OppNets is higher than the period T. Therefore, when a contact occurs, T_{ct} is uniformly distributed over T and also independent of random variable T_a. T_d is the lasting duration of the contact. We assume that the contact duration T_d is an i.i.d. stationary random variable with a cumulative distribution function (cdf) of $F_d(x)$. Let us define P_c (contact discovery probability) as the probability that a contact between two nodes can be discovered by one another in duty-cycle OppNets or the contact is an effective contact. There will be a set of different possibilities for calculating the contact discovery probability P_c, depending on the lengths of wake-up time T_{on}, period T, and contact duration T_d.

When $T_{\text{on}} < T_{\text{off}}$, to facilitate the calculation of the contact discovery probability P_c, we use the offset time T_a to divide the contact discovery probability P_c into three parts. Since T_a is uniformly distributed over the period T, we use P_{c1} to represent the contact discovery probability when $0 \le T_a \le T_{\text{on}}$, P_{c2} to represent the contact discovery probability when $T_{\text{on}} < T_a < T - T_{\text{on}}$, and P_{c3} to represent the contact discovery probability when $T - T_{\text{on}} \le T_a < T$. As shown in Fig. 1(a), nodes A and B have overlapped wake-up time with one another if the offset

time T_a is in the range of $[0, T_{\text{on}}]$. Under this situation, if T_{ct} is in the range of $(0, T_a)$ and (T_{on}, T), a contact happens when two nodes are not both in the wake-up state, and if T_{ct} is in the range of $[T_a, T_{\text{on}}]$, a contact happens when two nodes are both in the wake-up state. Since the effective contact contains two cases, then P_{c1} is the sum of three parts, which can be calculated as

$$\begin{aligned} P_{c1} = {} & \Pr\{0 \le T_a \le T_{\text{on}}\} \\ & \times [\Pr\{0 \le T_{\text{ct}} < T_a\} \Pr\{T_{\text{ct}} + T_d \ge T_a\} \\ & + \Pr\{T_a \le T_{\text{ct}} \le T_{\text{on}}\} \\ & + \Pr\{T_{\text{on}} < T_{\text{ct}} < T\} \Pr\{T_{\text{ct}} + T_d \ge T + T_a\}]. \end{aligned} \tag{2}$$

Note that two nodes do not have overlapped wake-up time with each other when offset time T_a is in the range of $(T_{\text{on}}, T - T_{\text{on}})$, which indicates that contacts cannot be discovered by each other under this situation. Therefore, P_{c2} is equal to 0 when $T_{\text{on}} < T_a < T - T_{\text{on}}$.

When $T - T_{\text{on}} \le T_a < T$, two nodes have an overlapped wake-up time with each other in the range of $[0, T_{\text{on}} + T_a - T]$. Under this situation, when T_{ct} is in the range of $[0, T_{\text{on}} + T_a - T]$, which indicates that a contact happens when two nodes are both in the wake-up state. When T_{ct} is in the range of $[T_{\text{on}} + T_a - T, T]$, which indicates a contact happens when two nodes are not both in the wake-up state. Since the effective contact contains two cases, similar to the calculating process of P_{c1}, then P_{c3} is the sum of two parts, which can be calculated as

$$\begin{aligned} P_{c3} = {} & \Pr\{T - T_{\text{on}} \le T_a \le T\} \\ & \times [\Pr\{0 \le T_{\text{ct}} \le T_{\text{on}} + T_a - T\} \\ & + \Pr\{T_{\text{on}} + T_a - T < T_{\text{ct}} \le T\} \\ & \times \Pr\{T_{\text{ct}} + T_d \ge T\}]. \end{aligned} \tag{3}$$

Therefore, the contact discovery probability P_c when $T_{\text{on}} < T_{\text{off}}$ can be calculated as $P_c = P_{c1} + P_{c2} + P_{c3}$.

When $T_{\text{on}} \ge T_{\text{off}}$, similar to the calculating process of P_c when $T_{\text{on}} < T_{\text{off}}$, we also use the offset time T_a to divide the contact discovery probability P_c into three parts. We use P'_{c1} to represent the contact discovery probability when $0 \le T_a \le T - T_{\text{on}}$, P'_{c2} to represent the contact discovery probability when $T - T_{\text{on}} < T_a < T_{\text{on}}$, and P'_{c3} to represent the contact discovery probability when $T_{\text{on}} \le T_a < T$. Then, we obtain the following expressions:

$$\begin{aligned} P'_{c1} = {} & \Pr\{0 \le T_a \le T - T_{\text{on}}\} \\ & \times [\Pr\{0 \le T_{\text{ct}} \le T_a\} \Pr\{T_{\text{ct}} + T_d \ge T_a\} \\ & + \Pr\{T_a < T_{\text{ct}} \le T_{\text{on}}\} + \Pr\{T_{\text{on}} < T_{\text{ct}} \le T\} \\ & \times \Pr\{T_{\text{ct}} + T_d \ge T + T_a\}] \end{aligned} \tag{4}$$

$$\begin{aligned} P'_{c2} = {} & \Pr\{T - T_{\text{on}} < T_a \le T_{\text{on}}\} \\ & \times [\Pr\{0 \le T_{\text{ct}} \le T_a - T + T_{\text{on}}\} \\ & + \Pr\{T_a - T + T_{\text{on}}) < T_{\text{ct}} \le T_a\} \\ & \times \Pr\{T_{\text{ct}} + T_d \ge T_a\} + \Pr\{T_a < T_{\text{ct}} \le T_{\text{on}}\} \\ & + \Pr\{T_{\text{on}} < T_{\text{ct}} \le T\} \Pr\{T_{\text{ct}} + T_d \ge T\}] \end{aligned} \tag{5}$$

$$\begin{aligned} P'_{c3} = {} & \Pr\{T_{\text{on}} < T_a < T\} \\ & \times [\Pr\{0 \le T_{\text{ct}} \le T_a - T + T_{\text{on}}\} \\ & + \Pr\{T_a - T + T_{\text{on}} < T_{\text{ct}} \le T\} \\ & \times \Pr\{T_{\text{ct}} + T_d \ge T\}]. \end{aligned} \tag{6}$$

Then, the contact discovery probability P_c when $T_{\text{on}} \ge T_{\text{off}}$ can be calculated as $P_c = P'_{c1} + P'_{c2} + P'_{c3}$.

B. Relationship Between Energy Consumption and the Contact Discovery Probability

Given the relationship among contact discovery probability P_c, period T, wake-up interval T_{on}, and contact duration T_d, we now explore the relationship between energy consumption and the contact discovery probability P_c under different situations. When contact duration T_d is distributed according to a given distribution, we can analytically obtain the relationship between energy consumption and the contact discovery probability. In [21], it was found that the cumulative contact duration in real mobility traces follows the power law distribution (Pareto distribution). Therefore, in this paper, we also assume that the contact duration T_d follows the power law distribution and can be expressed as

$$F_d(x) = \begin{cases} 0, & x < \tau \\ 1 - (x/\tau)^{-k}, & x \ge \tau \end{cases} \tag{7}$$

where τ is the minimum value of T_d, and k is the slope of the distribution.

Then, we get the expression of P_c under different situations by substituting (1) and (7) into (2)–(6). When $T_{\text{on}} < T_{\text{off}}$, P_{c1} and P_{c3} can be expressed as

$$P_{c1} = \frac{1}{T^2} \left\{ \frac{\tau^k \left[T^{2-k} - (T - T_{\text{on}})^{2-k}\right]}{(1-k)(2-k)} + 0.5T_{\text{on}}^2 + \tau T_{\text{on}} - \frac{\tau T_{\text{on}}}{1-k} \right\} \tag{8}$$

$$P_{c3} = \frac{1}{T^2} \left\{ \frac{\tau^k \left[T^{2-k} - (T - T_{\text{on}})^{2-k}\right]}{(1-k)(2-k)} + 0.5T_{\text{on}}^2 + \tau T_{\text{on}} - \frac{\tau T_{\text{on}}}{1-k} \right\}. \tag{9}$$

The Appendix describes how to obtain the given two expressions.

Since $P_{c2} = 0$, then P_c is the sum of the other two parts and can be expressed as

$$P_c = \frac{2}{T^2} \left\{ \frac{\tau^2 k \left[T^{2-k} - (T - T_{\text{on}})^{2-k}\right]}{(1-k)(2-k)} + 0.5T_{\text{on}}^2 + \tau T_{\text{on}} - \frac{\tau T_{\text{on}}}{1-k} \right\}. \tag{10}$$

When $T_{\text{on}} \ge T_{\text{off}}$, the expression of P_c is divided into three parts by τ: $0 < \tau \le T - T_{\text{on}}$, $T - T_{\text{on}} < \tau < 2(T - T_{\text{on}})$, and

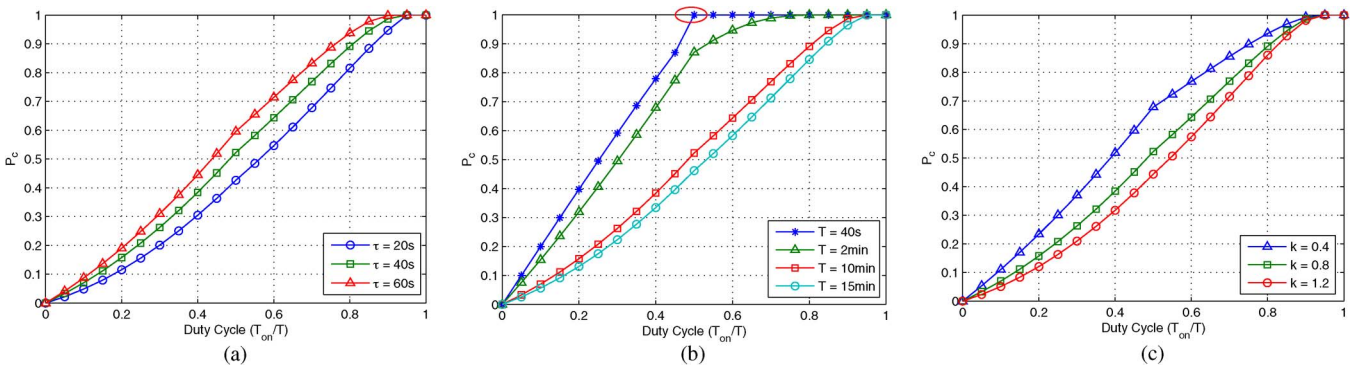


Fig. 2. Relationship between energy consumption and contact discovery probability under different situations. (a) Different τ. (b) Different T. (c) Different k.

$2(T - T_{\text{on}}) \leq \tau \leq T$. The computing process when $T_{\text{on}} \geq T_{\text{off}}$ is similar to the computing process when $T_{\text{on}} < T_{\text{off}}$; therefore, here, we omit the computing process. When $0 < \tau \leq T - T_{\text{on}}$, P_c is the sum of the three parts: P'_{c1}, P'_{c2}, and P'_{c3}. Then, we can obtain the following expression:

$$P_c = \frac{1}{T^2}\left\{\frac{2\tau^k\left[(2T-2T_{\text{on}})^{2-k} - (T-T_{\text{on}})^{2-k}\right]}{(1-k)(2-k)} + T_{\text{on}}^2 - \frac{2\tau(T-T_{\text{on}})}{1-k} + (4T_{\text{on}} - 2T) \times \left[\frac{(\tau^k(T-T_{\text{on}})^{1-k}}{1-k} - \frac{\tau}{1-k}\right] + 2\tau T_{\text{on}}\right\}. \quad (11)$$

When $T - T_{\text{on}} < \tau < 2(T - T_{\text{on}})$, similar to the approach given, we obtain the following expression:

$$P_c = \frac{1}{T^2}\left\{8TT_{\text{on}} - 4T_{\text{on}}^2 - 4\tau T_{\text{on}} + 4\tau T - \tau^2 - 3T^2 + \frac{2\tau^2\left[(2T-2T_{\text{on}})^{2-k} - \tau^{2-k}\right]}{(1-k)(2-k)} - \frac{2\tau(2T-2T_{\text{on}}-\tau)}{1-k}\right\}. \quad (12)$$

When $2(T - T_{\text{on}}) \leq \tau \leq T$ or $\tau > T$, similar to the approach above, we can obtain $P_c = 1$.

After obtaining the expression of P_c under different situations, here, we represent P_c graphically and show the relationship between energy consumption and contact discovery probability when the contact duration follows the power law distribution. Energy consumption is described as the duty cycle T_{on}/T, whereas a larger duty cycle to achieve a certain P_c means more energy consumption. Fig. 2 shows the relationship between energy consumption and contact discovery probability when the contact duration follows the power law distribution. It can be found that the contact discovery probability P_c is increasing along with the duty cycle, which means that nodes in the network have to consume more energy to increase the network performance. Moreover, when $T_{\text{on}} \geq T_{\text{off}}$ and $\tau \geq 2(T - T_{\text{on}})$, contact discovery probability P_c is always 100%, which means that P_c is always 100% when the duty cycle is no less than $\max\{0.5, 1 - (\tau/2T)\}$. In Fig. 2(a), it can be found that the contact discovery probability P_c is increasing along with τ, which means that larger τ needs less energy to achieve a certain P_c. In Fig. 2(b), it can be found that the contact discovery probability P_c increases when T decreases, and P_c reaches 100% more quickly when T is smaller; thus, smaller T needs less energy to achieve a certain P_c. However, T cannot be too small because small T means that the nodes have to switch between the wake-up state and the sleep state more often and that the switching process also consumes a lot of energy. Fig. 2(c) shows that contact discovery probability P_c increases when k decreases; thus, smaller k needs less energy to achieve a certain P_c.

This can be summarized with the following conclusions. Given that the contact duration T_d follows the power law distribution, we analytically obtain the relationship between energy consumption and contact discovery probability. It can be found that the contact discovery probability P_c increases when the duty cycle increases, and P_c is always 100% when the duty cycle is no less than $\max\{0.5, 1 - (\tau/2T)\}$. Fig. 2(b) shows the situation when $T = \tau = 40$ s. It can be found that P_c is always 100% when the duty cycle is no less than 50%. Moreover, the contact discovery probability P_c increases when τ increases, and decreases when k increases. From here, it can be found that the contact duration has a significant impact on the contact discovery probability. It is worth noting that the model proposed in [25] only gives the contact discovery probability when the contact duration T_d is a certain value; thus, results in [25] only analyze the relationship between energy consumption and the contact discovery probability when the contact duration T_d is a certain value.

V. Model Validation

Here, we use the real mobility trace *Infocom 06* collected from realistic environment to validate our proposed model. The *Infocom 06* trace was collected by 78 volunteers using iMotes with Bluetooth. Each device detects the vicinity every 120 s. When a device discovers other devices, it records the contact time and the ID of other devices. By using the information recorded by these mobile devices, we can obtain knowledge of the contact duration between two nodes. If a device is discovered in m contiguous scans, then the duration of the contact is the difference between the discovery time in the mth scan and the first scan. If a device is only detected in one scan, similar to the approach in [21], we treat the duration of contact as 120 s. Then, we look at the cumulative contact duration

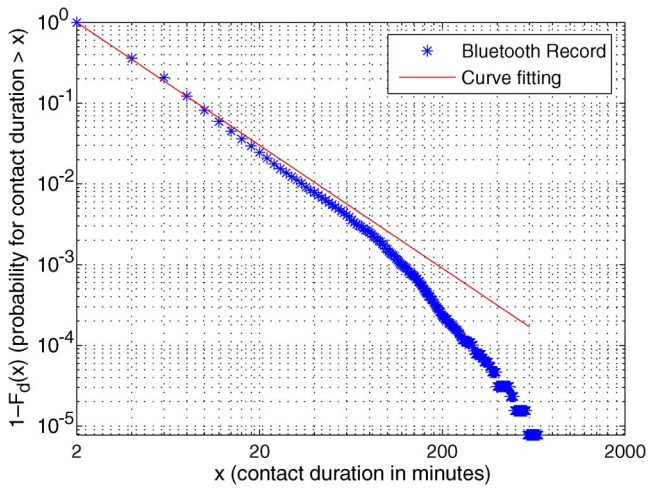


Fig. 3. Cumulative contact duration distribution.

distribution in the *Infocom 06* trace. Fig. 3 plots the $1 - F_d(x)$ curve in log–log scale. It can be found that the cumulative contact duration follows the power law distribution. By curve fitting, we can estimate $F_d(x) = 1 - (x/\tau)^{-k}$ with $\tau = 120$ s and $k = 1.523$. The fact that the cumulative contact duration follows the power law distribution has been also validated by other studies [21], [32].

After introducing the real mobility trace, we then use it to validate the correctness of our proposed model. Fig. 4 shows the comparison between the real-trace-driven simulation results and the theoretical results. Fig. 4(a) shows the comparison between the real-trace-driven simulation results and the theoretical results obtained in this paper. It can be found that the theoretical results are quite close to the real-trace-driven simulation results when the duty cycle is increasing, which confirms the correctness of our proposed model. Fig. 4(b) shows the comparison between the real-trace-driven simulation results and the theoretical results obtained in [25]. It can be found that the real-trace-driven simulation results and the theoretical results have large errors when the contact duration $T_d = 50$, 200, and 400 s, respectively. From here, it can be found that our proposed model is more suitable to the real environment than the model proposed in [25].

VI. Data Forwarding Strategy in Duty-Cycle OppNets

Here, we propose a data-forwarding strategy for duty-cycle OppNets. The proposed data-forwarding strategy aims to forward data copies toward the opportunistic forwarding paths, which maximize data delivery probability. To obtain the data delivery probability of each opportunistic forwarding path in duty-cycle OppNets, we start off by analyzing the pairwise contact duration distribution in real mobility traces.

A. Pairwise Contact Duration Distribution in Real Mobility Traces

In Section IV, we have given the expression of the contact discovery probability P_c when the contact duration follows the

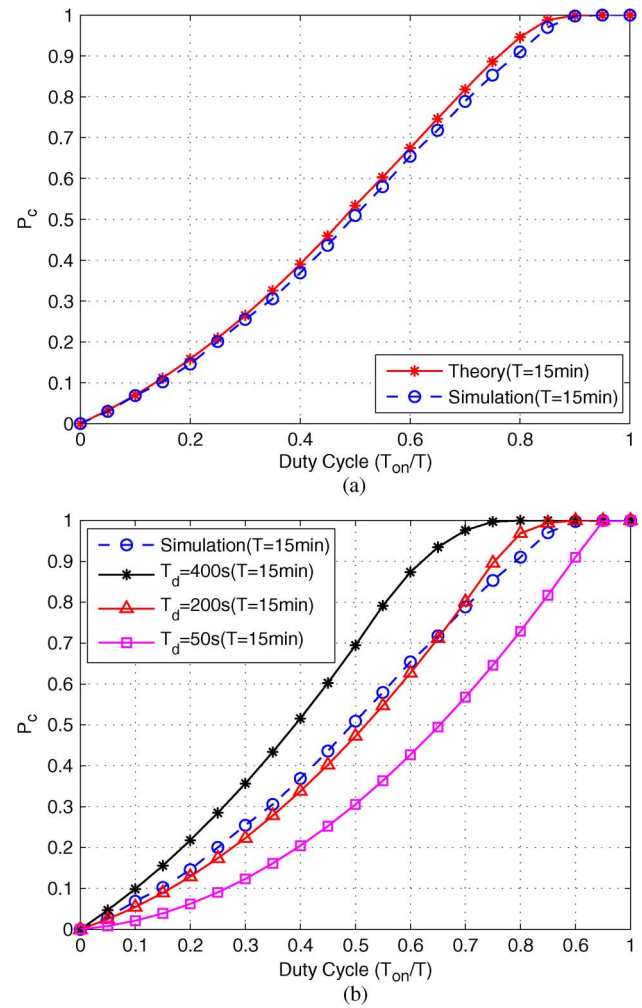


Fig. 4. Comparison between real-trace-driven simulation results and theoretical results. (a) Validation of our proposed model. (b) Validation of the model proposed in [25].

power law distribution. Note that the previous studies [21], [32] only show that the cumulative contact duration follows the power law distribution, and they do not take into account the distribution of the pairwise contact duration. Although some studies [33], [34] have the assumption that the pairwise contact duration follows the power law distribution, they did not validate it experimentally. Therefore, in this paper, we first use two real mobility traces, i.e., *Infocom 06* and *MIT Reality,* to validate the hypothesis that the pairwise contact duration also follows the power law distribution.

To validate this hypothesis, we conduct chi-square hypothesis test [35] on each contacted node pair in the given traces, to test whether the hypothesis "the pairwise contact duration also follows the power law distribution." Similar to the approach in Section V, we first use curve fitting to estimate the pairwise contact duration distribution parameters k and τ of each contacted node pair, where k is the slope of the distribution, and τ is the minimum value of the contact duration. Then, we conduct a chi-square hypothesis test on each contacted node pair in the given traces by comparing the contact duration sample frequencies with the theoretical probabilities. Since the power law distribution is continuous, in the chi-square tests, we divide the

TABLE I
ACCEPTANCE RATIO OF CHI-SQUARE TESTS FOR THE *Infocom 06* TRACE

Infocom06	5	10	15	20	25
0.95	0.8004	0.8897	0.9183	0.9324	0.941
0.75	0.9512	0.9682	0.9734	0.9754	0.9772
0.5	0.9855	0.987	0.9872	0.9871	0.9843

TABLE II
ACCEPTANCE RATIO OF CHI-SQUARE TESTS FOR THE *MIT Reality* TRACE

MIT Reality	5	10	15	20	25
0.95	0.831	0.9095	0.9371	0.9508	0.9586
0.75	0.9449	0.9693	0.978	0.9818	0.9839
0.5	0.9839	0.9904	0.9924	0.993	0.993

range of the sample values into several test intervals and compare the sample frequencies with theoretical probabilities on each interval. The results of the acceptance ratio for the given traces under different significance levels α are listed in Tables I and II. The results show that more than 80% of the contacted node pairs in the given traces pass the test when the test interval is 5, and more than 90% of the contacted node pairs in the given traces pass the test when the test interval increases to 15.

Based on the given experimental results, we validate the hypothesis that the pairwise contact duration follows the power law distribution. Therefore, the pairwise contact discovery probability $P_c(ij)$ between nodes i and j can be expressed as equations in Section IV when the contact duration follows the power law distribution. Then, we can use this pairwise contact discovery probability to investigate node contact patterns in duty-cycle OppNets.

B. Opportunistic Forwarding Path in Duty-Cycle OppNets

In OppNets, nodes' contact can be described as network-connected graph $G(V, E)$. The random contact process between nodes i and j can be modeled as $e_{ij} \in E$, where $i, j \in V$. Some recent studies [3], [36], [37] have found that the pairwise intercontact time in real mobility traces follows the exponential distribution. Specifically, in [3], a chi-square hypothesis test is conducted on each contacted node pair in the *Infocom 06* and *MIT Reality* traces to test whether "the pairwise node intercontact time follows the exponential distribution." Their results demonstrate that, when enough number of test intervals ($\geq$ 10) is used, over 85% of the contacted node pairs in the given traces pass the test. Therefore, in this paper, we also assume that the pairwise intercontact time in OppNets follows the exponential distribution. Then, the contact frequency λ_{ij} between nodes i and j is indicated by the contact rate and can be computed by the following time average method:

$$\lambda_{ij} = \frac{n}{\Sigma_{l=1}^{n} T_{ij}^{l}}. \tag{13}$$

Thus, the pdf of the intercontact time X_{ij} between nodes i and j can be expressed as

$$f_{X_{ij}}(t) = \lambda_{ij} e^{-\lambda_{ij} t}. \tag{14}$$

Note that, in duty-cycle OppNets, each contact between two nodes has a pairwise contact discovery probability, which is relevant to the contact duration between two nodes. Therefore, according to the thinning property of the Poisson process [38], the pairwise intercontact time X'_{ij} between nodes i and j in duty-cycle OppNets also follows the exponential distribution, and the contact frequency λ'_{ij} between nodes i and j in duty-cycle OppNets can be calculated as $P_c(ij)\lambda_{ij}$, where $P_c(ij)$ is the pairwise contact discovery probability between nodes i and j.

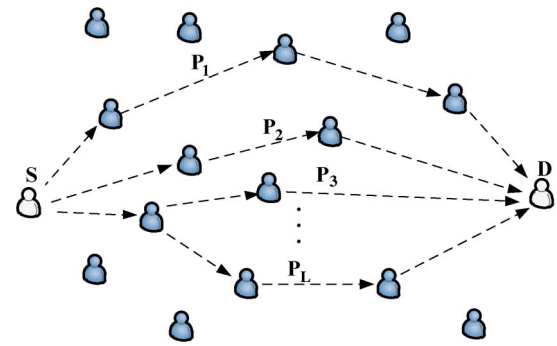


Fig. 5. Opportunistic forwarding path in OppNets.

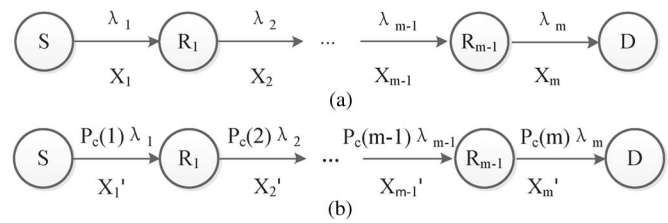


Fig. 6. Edge weights of a certain opportunistic forwarding path between S and D (a) Path edge weights in OppNets. (b) Path edge weights in duty-cycle OppNets.

Thus, the pdf of the pairwise intercontact time X'_{ij} between nodes i and j in duty-cycle OppNets can be expressed as

$$f_{X'_{ij}}(t) = P_c(ij)\lambda_{ij} e^{-P_c(ij)\lambda_{ij} t}. \tag{15}$$

After obtaining the distribution of the pairwise intercontact time in duty-cycle OppNets, we then introduce the concept of the opportunistic forwarding path [39]. As shown in Fig. 5, we assume that there is a certain data item that is transferred from the source node S to the destination node D, and there are L opportunistic forwarding paths from S to D. As shown in Fig. 6, the definition of the opportunistic forwarding path is given as follows.

Definition 1: In OppNets, a certain m-hop opportunistic forwarding path between S and D, which is denoted as l, consists of node set $\{S, R_1, R_2, \ldots, R_{m-1}, D\}$ and edge set $\{e_1, e_2, \ldots, e_m\}$ with edge weights $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, where $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ are the contact rates (or contact frequencies) of each adjacent node pair along the opportunistic forwarding path in OppNets. In duty-cycle OppNets, since each contact between two nodes has a pairwise contact discovery probability, the corresponding edge weights change to $\{P_c(1)\lambda_1, P_c(2)\lambda_2, \ldots, P_c(m)\lambda_m\}$.

It is worth noticing that, when m is equal to 1, which means that there exists a direct edge between S and D, the node set will be changed to $\{S, D\}$, and the corresponding edge weight

is λ_{SD}. Note that the intercontact time X_i' between two nodes R_{i-1} and R_i in duty-cycle OppNets follows the exponential distribution with a pdf of $f_{X_i'}(t) = P_c(i)\lambda_i e^{-P_c(i)\lambda_i t}$. As a result, the total time to transfer a data item from S to D along opportunistic forwarding path l in duty-cycle OppNets is $Y_l = \sum_{i=1}^{m} X_i'$, and the pdf $f_{Y_l}(t)$ can be calculated as

$$f_{Y_l}(t) = f_{X_1'}(t) \otimes f_{X_2'}(t), \ldots, \otimes f_{X_m'}(t) \tag{16}$$

where $\otimes$ is the convolution operator.

Then, by referring to the theoretical results in [3], we have the following theorem.

Theorem 1: For a certain m-hop opportunistic forwarding path l with edge weights $\{P_c(1)\lambda_1, P_c(2)\lambda_2, \ldots, P_c(m)\lambda_m\}$, when m is larger than 1, $p_{Y_l}(t)$ is expressed as

$$f_{Y_l}(t) = \sum_{i=1}^{m} C_i^m f_{X_i'}(t) \tag{17}$$

where the coefficients are given as follows:

$$C_i^m = \prod_{j=1, j\neq i}^{m} \frac{P_c(j)\lambda_j}{P_c(j)\lambda_j - P_c(i)\lambda_i}. \tag{18}$$

From $f_{Y_l}(t)$, the probability that a certain data item is successfully delivered from S to D within time T_0 along an opportunistic forwarding path l in duty-cycle OppNets is expressed as

$$\begin{aligned}\Pr_{Y_l}(T_0) = P(Y_l < T_0) &= \int_0^{T_0} f_{Y_l}(t)dt \\ &= \sum_{i=1}^{m} C_i^m (1 - e^{-P_c(i)\lambda_i T_0}).\end{aligned} \tag{19}$$

Note that if there exists an edge directly between nodes S and D, which means that m is equal to 1, then the probability that a certain data item is successfully delivered from S to D within time T_0 in duty-cycle OppNets is

$$\begin{aligned}\Pr_{Y_l}(T_0) = P(Y_l < T_0) &= \int_0^{T_0} P_c(SD)\lambda_{SD} e^{-P_c(SD)\lambda_{SD} t} dt \\ &= 1 - e^{-P_c(SD)\lambda_{SD} T_0}\end{aligned} \tag{20}$$

where $P_c(SD)$ is the pairwise contact discovery probability between nodes S and D.

C. Maximum Data Delivery Probability Forwarding Strategy

After obtaining the data delivery probability along a certain opportunistic forwarding path in duty-cycle OppNets, here, we introduce our proposed data-forwarding strategy for duty-cycle OppNets. Take two nodes A and B as an example. When node A encounters node B, we assume that node A has a copy of the data item that is delivered from S to D, and the remaining time-to-live (TTL) is T_r; then, it has to decide whether or not to forward a data copy to node B. Note that there are L_A opportunistic forwarding paths between A and D, L_B opportunistic forwarding paths between B and D, and each opportunistic forwarding path has a data delivery probability. Therefore, we use $\Pr_{\max}^{AD}(T_r)$ and $\Pr_{\max}^{BD}(T_r)$ to indicate the maximum data delivery probability of these L_A and L_B opportunistic forwarding paths in duty-cycle OppNets, respectively, which are calculated as

$$\Pr_{\max}^{AD}(T_r) = \max\left\{\Pr_{Y_1}^{AD}(T_r), \Pr_{Y_2}^{AD}(T_r), \ldots, \Pr_{Y_{L_A}}^{AD}(T_r)\right\} \tag{21}$$

$$\Pr_{\max}^{BD}(T_r) = \max\left\{\Pr_{Y_1}^{BD}(T_r), \Pr_{Y_2}^{BD}(T_r), \ldots, \Pr_{Y_{L_B}}^{BD}(T_r)\right\}. \tag{22}$$

Then, we use the maximum data delivery probability as the data-forwarding metrics and introduce our proposed data-forwarding strategy. The pseudocode in Algorithm 1 depicts the basic operations of our proposed data-forwarding strategy from a node's perspective. When node A encounters node B, they first compare each data's maximum data delivery probability in their buffer with each other. Then, the node with a lower maximum data delivery probability forwards a copy of the corresponding data to the node with a higher maximum data delivery probability.

Algorithm 1 Maximum data delivery probability forwarding strategy

```
When node A encounters node B and the remaining TTL is T_r
for all data(i) stored at the buffer of node A
  if node B does not have data(i) in its buffer
    if data(i) · destination = B or data(i) · Pr^BD_max(T_r) > data(i) · Pr^AD_max(T_r) then
      A forwards a copy of data(i) to B
    end if
  end if
end for
node B do the same loop as node A
```

D. Overhead Reduction

As introduced earlier, to conduct the Algorithm 1, nodes A and B need to calculate the maximum data delivery probability among L_A opportunistic forwarding paths between A and D, as well as L_B opportunistic forwarding paths between B and D using (21) and (22), respectively. If the length of these opportunistic forwarding paths is long or the value of L_A and L_B is large, the calculation overhead for nodes A and B will be huge. Moreover, the overhead for nodes A and B to collect node contact patterns (including the pairwise intercontact time and the pairwise contact duration) of these L_A and L_B opportunistic forwarding paths will be also huge.

To reduce the calculation and collection overhead, in this paper, we exclude those opportunistic forwarding paths, which are larger than three hops. This is because most node pairs in the *Infocom 06* and *MIT Reality* traces only need less than three hops to connect each other. As shown in Fig. 7, when the number of minimum hops to connect node pairs in the *Infocom 06* trace is less than three, the cumulative percentage is more than

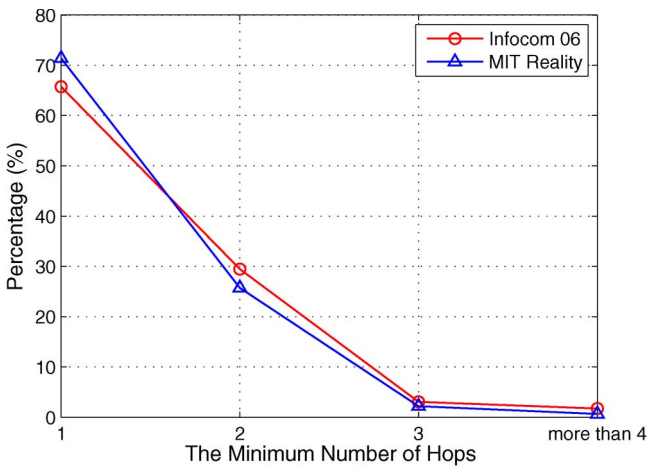


Fig. 7. Percentage of the number of minimum hops to connect nodes pairs in the *Infocom 06* and *MIT Reality* traces.

95%, whereas the corresponding value in the *MIT Reality* trace is more than 97%. Since a larger number of hops means less data delivery probability of the opportunistic forwarding path, we can exclude those opportunistic forwarding paths, which are larger than three hops. Furthermore, in the calculation process of the maximum data delivery probability, we also exclude those opportunistic forwarding paths, which are subtrees of a certain opportunistic forwarding path. For example, we assume that there are three opportunistic forwarding paths from S to D, i.e., path 1 $S{-}A{-}D$, path 2 $S{-}A{-}B{-}D$ and path 3 $S{-}A{-}B{-}C{-}D$. We can exclude paths 2 and 3. This is because $\mathrm{Pr}_1(T_r)$ is absolutely larger than $\mathrm{Pr}_2(T_r)$ and $\mathrm{Pr}_3(T_r)$.

VII. Performance Evaluation

Here, we mainly focus on evaluating the performance of our proposed data-forwarding strategy, and investigating the impact of some parameters on the performance of our proposed data-forwarding strategy. Here, we use *Maximum* to indicate our proposed data-forwarding strategy.

A. Simulation Setup

We evaluate the performance of our proposed data-forwarding strategy *Maximum* in terms of delivery ratio and delivery cost. The delivery ratio is the ratio of data successfully delivered by nodes, and the delivery cost is the average number of data copies forwarded in the network. The delivery delay is not taken into consideration, as long as the data can be delivered on time. In our simulation studies, we compare our proposed data-forwarding strategy *Maximum* with the following three data-forwarding strategies.

1) **Epidemic Routing**: Data copies are simply flooded to nodes in network.
2) **Bubble Rap**: Data copies are first forwarded to nodes that have higher global centrality. When data copies are forwarded to the same community as the destination node, then local centrality will be used instead of global centrality as the forwarding metrics, and data copies continue to bubble up until the destination is reached or the data expire.
3) **Prophet**: Nodes in the network use the past contact history to predict the probability of meeting a node again, and data copies are forwarded to nodes that have higher contact probability for the destination node.

TABLE III
Basic Statistics of the Traces

Trace	*Infocom 06*	*MIT Reality*
Device	iMote	Smart Phones
Network type	Bluetooth	Bluetooth
Duration (days)	3	246
No. of internal contacts	182,951	114,046
Granularity (seconds)	120	300
No. of devices	78	97
Contact frequency/pair/day	6.7	0.024

We use two real mobility traces *Infocom 06* [30] and *MIT Reality* [31] collected from real environments to evaluate the performance of the selected data-forwarding strategies. Users in these two traces are all carrying Bluetooth-enabled portable devices, which record contacts by periodically detecting their peers nearby. The traces cover various types of corporate environments and have various experiment periods. The details of the traces are summarized in Table III.

We use a part of the traces (the first day of *Infocom 06*, and the September and October of *MIT Reality*) to model and characterize node contact patterns (i.e., calculate the contact frequency λ_{ij} and distribution parameters of the pairwise contact duration for our proposed data-forwarding strategy, and calculate the betweenness centrality and form a community for Bubble Rap, respectively). We use another part of the traces (the second day of *Infocom 06* and the November of *MIT Reality*) to evaluate the performance of the selected data-forwarding strategies.

B. Performance Comparison

Here, we compare the performance of our proposed data-forwarding strategy with other existing data-forwarding strategies in the *Infocom 06* and *MIT Reality* traces, respectively. Here, Epidemic Routing represents the baseline for the best delivery ratio performance and the baseline for the worst delivery cost performance. This is because Epidemic Routing always finds the best possible opportunistic forwarding path to the destination but is expensive in terms of delivery cost since data copies in Epidemic Routing are simply flooded to nodes in the network. As a result, the ultimate goal of our proposed data-forwarding strategy *Maximum* is to achieve delivery ratio as close to Epidemic Routing as possible and achieve delivery cost as small as possible.

Fig. 8 shows the performance comparison of our proposed data-forwarding strategy *Maximum* with other existing data-forwarding strategies when T is 10 min and TTL is 1 h in the *Infocom 06* trace. It can be found that the delivery ratio and cost are both tightly related to the duty cycle (T_{on}/T). As the duty cycle increases from 10% to 90%, the delivery ratio and cost both increase, particularly when the duty cycle is less than 50%. This is because fewer contacts will be missed when the duty

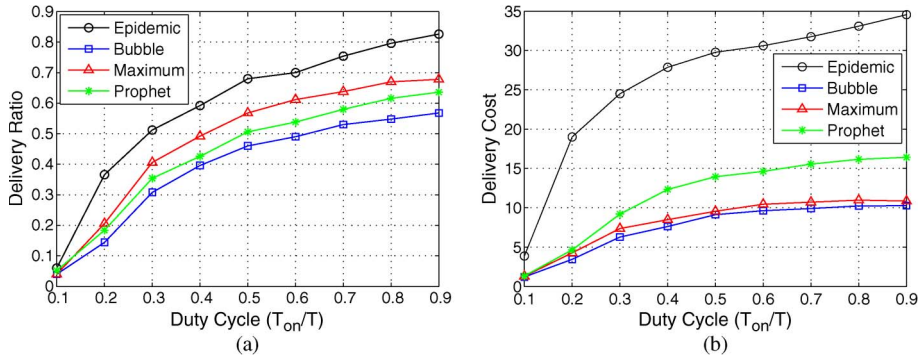


Fig. 8. Performance comparison of *Maximum* with other existing data-forwarding strategies when T is 10 min and TTL is 1 h in the *Infocom 06* trace. (a) Delivery ratio. (b) Delivery cost.

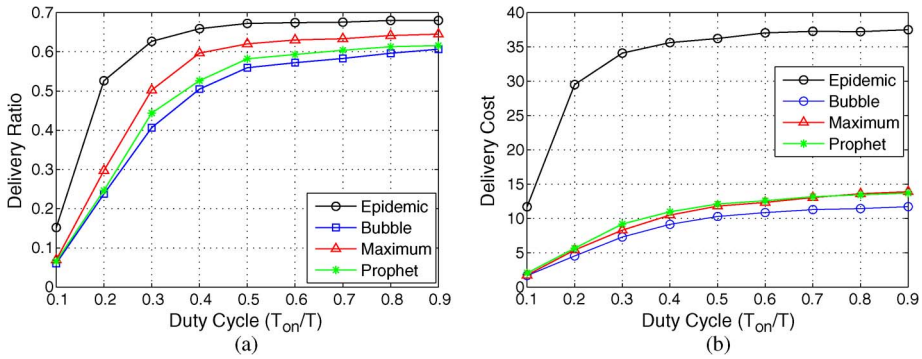


Fig. 9. Performance comparison of *Maximum* with other existing data-forwarding strategies when T is 20 min and TTL is five days in the *MIT Reality* trace. (a) Delivery ratio. (b) Delivery cost.

cycle increases, or more contacts can be used for data exchanges when the duty cycle increases, resulting in the increase in the delivery ratio and cost. Moreover, as the duty-cycle increases, Epidemic Routing performs best in terms of delivery ratio and performs worst in terms of delivery cost, as expected. Although Epidemic Routing outperforms *Maximum* in terms of delivery ratio, however, the delivery cost of Epidemic Routing is almost three to five times of *Maximum*, whereas the delivery ratio of Epidemic Routing is only 10%–15% larger than *Maximum*. *Maximum* outperforms Bubble Rap and Prophet in terms of delivery ratio, and the delivery cost is only slightly larger than that of Bubble Rap. The main reason is that *Maximum* takes node contact patterns in duty-cycle OppNets into consideration and manages to forward data copies toward a better path to the destination in duty-cycle OppNets. Therefore, *Maximum* can achieve a high delivery ratio with a low delivery cost. It is worth noting that Bubble Rap performs worst in terms of delivery ratio in duty-cycle OppNets, although the delivery cost of Bubble is the least. The main reason is that Bubble Rap is a node contact pattern-based data-forwarding strategy that is designed in OppNets; however, the node contact pattern in duty-cycle OppNets is obviously different from that in OppNets. Therefore, the social relationship between nodes and the centrality of nodes in OppNets is not suitable for those in duty-cycle OppNets, which makes it difficult to forward data copies toward a better path to the destination in duty-cycle OppNets.

Fig. 9 shows the performance comparison of our proposed data-forwarding strategy *Maximum* with other existing data-forwarding strategies when T is 20 min and TTL is five days in the *MIT Reality* trace. It can be found that, similar to the results in Fig. 8, the delivery ratio and cost are also both tightly related to the duty cycle. Moreover, as the duty cycle increases, Epidemic Routing also performs best in terms of delivery ratio and performs worst in terms of delivery cost as expected. In addition, our proposed data-forwarding strategy outperforms Bubble Rap and Prophet in terms of delivery ratio, and the delivery cost is only slightly larger than that of Bubble Rap. It is worth noting that the delivery cost of Prophet severely decreases in Fig. 9, compared with the results in Fig. 8. The main reason is that the contacts in the *MIT Reality* trace are much sparser than that in the *Infocom 06* trace, particularly when the duty-cycle operation is applied to the network. Therefore, it is difficult to compare nodes' probabilities of contacting the destination to make forwarding decisions, resulting in the decrease in delivery cost in Prophet.

To summarize, the delivery ratio and cost are both tightly related to the duty cycle in the *Infocom 06* and *MIT Reality* traces. Moreover, *Maximum* outperforms Bubble Rap and Prophet in terms of delivery ratio with reasonable delivery cost. Therefore, compared with other existing data-forwarding strategies, *Maximum* is more efficient and suitable for duty-cycle OppNets.

C. Impact of T and TTL

Here, we carry out experiments when T is different in the *Infocom 06* trace, and when TTL is different in the *MIT*

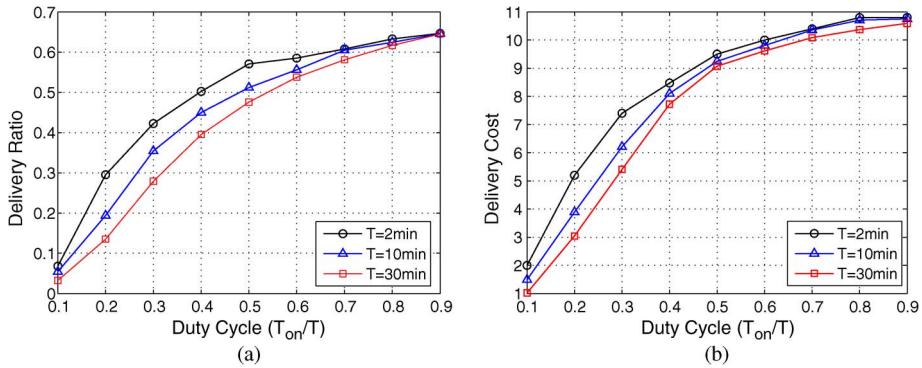


Fig. 10. Performance of *Maximum* when T is different in the *Infocom 06* trace. (a) Delivery ratio. (b) Delivery cost.

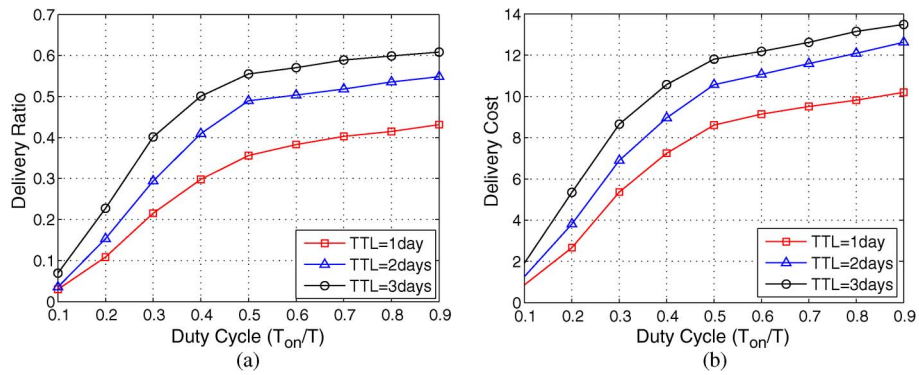


Fig. 11. Performance of *Maximum* when TTL is different in the *MIT Reality* trace. (a) Delivery ratio. (b) Delivery cost.

Reality trace, aiming to check the impact of the changing parameters on the performance of our proposed data-forwarding strategy.

Fig. 10 shows the impact of T on the performance of our proposed data-forwarding strategy in the *Infocom 06* trace. As T increases from 2 to 30 min, the delivery ratio and cost of our proposed data-forwarding strategy both decrease. The main reason is that the contact discovery probability decreases as T increases. Therefore, more contacts will be missed when T increases, or fewer contacts can be used for data exchanging, which causes the decrease in delivery ratio and cost.

Fig. 11 shows the impact of TTL on the performance of our proposed data-forwarding strategy in the *MIT Reality* trace. It can be found that the delivery ratio and cost of our proposed data-forwarding strategy both increase as TTL increases from one to three days. A reasonable explanation is that, when TTL increases, nodes have more time to deliver the data to the destination. Therefore, more nodes are involved in the data-forwarding process, resulting in the increase in the delivery ratio and cost.

In summary, T and TTL both have a significant impact on the performance of our proposed data-forwarding strategy. Although increasing the value of T can decrease the delivery cost, it also decreases the delivery ratio of our proposed data-forwarding strategy. Similarly, increasing the value of TTL can increase the delivery ratio, but it also obviously increases the delivery cost of our proposed data-forwarding strategy. Therefore, we should choose an appropriate value of T and TTL according to different applications.

VIII. Conclusion

In this paper, we have proposed a model to investigate the contact process in duty-cycle OppNets and analyzed the relationship between energy consumption and the contact discovery probability under different situations. Moreover, we used real-trace-driven simulations to validate the correctness of our proposed model. Then, based on the proposed model, we have proposed a novel data-forwarding strategy for duty-cycle OppNets. The proposed data-forwarding strategy uses node contact patterns to aid the design of the data-forwarding strategy for duty-cycle OppNets and manages to forward data copies toward the opportunistic forwarding paths that maximize data delivery probability. Extensive real-trace-driven simulation results show that our proposed data-forwarding strategy is close to Epidemic Routing in terms of delivery ratio but with significantly reduced delivery cost. Additionally, our proposed data-forwarding strategy outperforms Bubble Rap and Prophet in terms of delivery ratio with reasonable delivery cost.

Appendix

When $T_{\text{on}} < T_{\text{off}}$, since the contact duration T_d follows the power law distribution, as shown (7), then the expression of P_c is divided into three parts by $\tau : 0 < \tau \leq T_{\text{on}}$, $T_{\text{on}} < \tau \leq T - T_{\text{on}}$, and $T - T_{\text{on}} < \tau \leq T$. When $0 < \tau \leq T_{\text{on}}$, according to (2), P_{c1} can be calculated as

$$P_{c1} = \frac{1}{T^2} \int_0^{T_{\text{on}}} dT_a \left(\int_0^{T_a} \Pr\{T_{\text{ct}} + T_d \geq T_a\} dT_{\text{ct}} + \int_{T_a}^{T_{\text{on}}} dT_{\text{ct}} \right.$$

$$
\begin{aligned}
&\quad + \int_{T_{\text{on}}}^{T} \Pr\{T_{\text{ct}} + T_d \geq T + T_a\} dT_{\text{ct}} \Bigg) \\
&= \frac{1}{T^2} \Bigg\{ \int_0^{\tau} dT_a \Bigg[\int_0^{T_a} dt + T_{\text{on}} - T_a + \int_{T_a}^{\tau} dt \\
&\qquad + \int_{\tau}^{T+T_a-T_{\text{on}}} \left(\frac{t}{\tau}\right)^{-k} dt \Bigg] \\
&\quad + \int_{\tau}^{T_{\text{on}}} dT_a \Bigg[\int_0^{\tau} dt + \int_{\tau}^{T_a} \left(\frac{t}{\tau}\right)^{-k} dt + T_{\text{on}} - T_a \\
&\qquad + \int_{T_a}^{T+T_a-T_{\text{on}}} \left(\frac{t}{\tau}\right)^{-k} dt \Bigg] \Bigg\} \\
&= \frac{1}{T^2} \Bigg\{ \frac{\tau^k \left[T^{2-k} - (T - T_{\text{on}})^{2-k}\right]}{(1-k)(2-k)} \\
&\qquad + 0.5 T_{\text{on}}^2 - \frac{\tau T_{\text{on}}}{1-k} + \tau T_{\text{on}} \Bigg\}. \qquad (23)
\end{aligned}
$$

Similar to the calculating process given, according to (3), we can obtain that $P_{c3} = P_{c1}$.

When $T_{\text{on}} < \tau \leq T - T_{\text{on}}$ and $T - T_{\text{on}} < \tau \leq T$, similar to the calculating process given and according to (2) and (3), we can obtain that the expressions of P_{c1} and P_{c3} are the same as the expression when $0 < \tau \leq T_{\text{on}}$. Thus, P_c can be expressed as (10).

REFERENCES

[1] F. Li and J. Wu, "MOPS: Providing content-based service in disruption-tolerant networks," in *Proc. IEEE ICDCS*, 2009, pp. 526–533.

[2] J. Wu and Y. Wang, "Social feature-based multi-path routing in delay tolerant networks," in *Proc. IEEE INFOCOM*, 2012, pp. 1368–1376.

[3] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: A social network perspective," in *Proc. ACM MobiHoc*, 2009, pp. 299–308.

[4] H. Zhou, J. Chen, J. Fan, Y. Du, and S. K. Das, "Consub: Incentive-based content subscribing in selfish opportunistic mobile networks," *IEEE J. Sel. Areas Commun.*, 2013, DOI: 10.1109/JSAC.2013.SUP.0513058.

[5] J. Fan, J. Chen, Y. Du, P. Wang, and Y. Sun, "Delque: A socially-aware delegation query scheme in delay tolerant networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2181–2193, Jun. 2011.

[6] J. Fan, J. Chen, Y. Du, W. Gao, J. Wu, and Y. Sun, "Geocommunity-based broadcasting for data dissemination in mobile social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 4, pp. 734–743, Apr. 2013.

[7] W. Rao, K. Zhao, Y. Zhang, P. Hui, and S. Tarkoma, "Maximizing timely content advertising in DTNs," in *Proc. IEEE SECON*, 2012, pp. 254–262.

[8] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," in *Proc. IEEE INFOCOM*, 2001, pp. 1548–1557.

[9] M. Stemm and R. H. Katz, "Measuring and reducing energy consumption of network interfaces in hand-held devices," *IEICE Trans. Commun.*, vol. 80, no. 8, pp. 1125–1131, Aug. 1997.

[10] E. Shih, P. Bahl, and M. J. Sinclair, "Wake on wireless: An event driven energy saving strategy for battery operated devices," in *Proc. ACM MobiCom*, 2002, pp. 160–171.

[11] N. Banerjee, M. D. Corner, and B. N. Levine, "Design and field experimentation of an energy-efficient architecture for dtn throwboxes," *IEEE/ACM Trans. Netw.*, vol. 18, no. 2, pp. 554–567, Apr. 2010.

[12] Y. Sun, O. Gurewitz, S. Du, L. Tang, and D. B. Johnson, "ADB: An efficient multihop broadcast protocol based on asynchronous duty-cycling in wireless sensor networks," in *Proc. ACM SenSys*, 2009, pp. 43–56.

[13] Z. Li, M. Li, and Y. Liu, "Towards energy-fairness in asynchronous duty-cycling sensor networks," in *Proc. IEEE INFOCOM*, 2012, pp. 801–809.

[14] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli, "Age matters: Efficient route discovery in mobile ad hoc networks using encounter ages," in *Proc. ACM MobiHoc*, 2003, pp. 257–266.

[15] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *Lecture Notes Comput. Sci.*, vol. 3126, pp. 239–254, 2004.

[16] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation forwarding," in *Proc. ACM MobiHoc*, 2008, pp. 251–260.

[17] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 606–620, Jun. 2007.

[18] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay tolerant networks," in *Proc. ACM MobiHoc*, 2008, pp. 241–250.

[19] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant manets," in *Proc. ACM MobiHoc*, 2007, pp. 32–40.

[20] P. Hui and J. Crowcroft, "How small labels create big improvements," in *Proc. IEEE Workshop Intermittently Connect. Mobile Ad Hoc Netw.*, 2007, pp. 65–70.

[21] W. Wang, V. Srinivasan, and M. Motani, "Adaptive contact probing mechanisms for delay tolerant applications," in *Proc. ACM MobiCom*, 2007, pp. 230–241.

[22] S. Qin, G. Feng, and Y. Zhang, "How the contact-probing mechanism affects the transmission capacity of delay-tolerant networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1825–1834, May 2011.

[23] C. Drula, C. Amza, F. Rousseau, and A. Duda, "Adaptive energy conserving algorithms for neighbor discovery in opportunistic bluetooth networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 1, pp. 96–107, Jan. 2007.

[24] H. Zhou, H. Zheng, J. Wu, and J. Chen, "Energy-efficient contact probing in opportunistic mobile networks," in *Proc. ICCCN*, 2013, pp. 1–7.

[25] O. Trullols-Cruces, J. Morillo-Pozo, J. M. Barcelo-Ordinas, and J. Garcia-Vidal, "Power saving trade-offs in delay/disruptive tolerant networks," in *Proc. IEEE WoWMoM*, 2011, pp. 1–9.

[26] H. Zhou, H. Zhao, and J. Chen, "Energy saving and network connectivity tradeoff in opportunistic mobile networks," in *Proc. IEEE GLOBECOM*, 2012, pp. 524–529.

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

[28] Y. C. Tseng, S. Y. Ni, Y. S. Chen, and J. P. Sheu, "The broadcast storm problem in a mobile ad hoc network," *Wireless Netw.*, vol. 8, no. 2/3, pp. 153–167, Mar.–May 2002.

[29] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: An efficient routing scheme for intermittently connected mobile networks," in *Proc. ACM SIGCOMM Workshop Delay-Tolerant Netw.*, 2005, pp. 252–259.

[30] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, CRAWDAD data set cambridge/haggle (v. 2009-05-29), 2009. [Online]. Available: http://crawdad.cs.dartmouth.edu/cambridge/haggle

[31] N. Eagle, A. Pentland, and D. Lazer, "Inferring social network structure using mobile phone data," *Proc. Nat. Acad. Sci.*, vol. 106, no. 36, pp. 15 274–15 278, Sep. 2009.

[32] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Pocket switched networks: Real-world mobility and its consequences for opportunistic," Univ. Cambridge, Comput. Lab., Cambridge, U.K., Tech. Rep. UCAM-CL-TR-617, 2005.

[33] X. Zhuo, Q. Li, G. Cao, Y. Dai, B. Szymanski, and T. La Porta, "Social-based cooperative caching in dtns: A contact duration aware approach," in *Proc. IEEE MASS*, 2011, pp. 92–101.

[34] X. Zhuo, Q. Li, W. Gao, G. Cao, and Y. Dai, "Contact duration aware data replication in delay tolerant networks," in *Proc. IEEE ICNP*, 2011, pp. 236–245.

[35] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*. New York, NY, USA: Wiley-Interscience, 1996.

[36] H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L. M. Ni, "Recognizing exponential inter-contact time in VANETs," in *Proc. IEEE INFOCOM*, 2010, pp. 1–5.

[37] V. Conan, J. Leguay, and T. Friedman, "Characterizing pairwise inter-contact patterns in delay tolerant networks," in *Proc. ACM SenSys*, 2007, pp. 321–334.

[38] J. Grandell, *Mixed Poisson Process*. Boston, MA, USA: Chapman & Hall, 1997.

[39] S. Prabhavat, H. Nishiyama, N. Ansari, and N. Kato, "On load distribution over multipath networks," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 3, pp. 662–680, Third Quar., 2012.

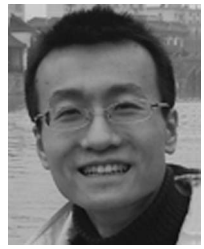
Huan Zhou is currently working toward the Ph.D. degree with the Department of Control Science and Engineering, State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China.

From November 2012 to May 2013, he was a Visiting Scholar with Temple University, Philadelphia, PA, USA. He is currently a member of the Networked Sensing and Control Group, Key Laboratory of Industrial Control Technology, Zhejiang University. His research interests include data transmission and energy saving in mobile social networks and opportunistic mobile networks.

Jiming Chen (M'08–SM'11) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2000 and 2005, respectively.

In 2006, he was a Visiting Researcher with the National Institute for Research in Computer Science and Control (INRIA), Paris, France; in 2007, with the National University of Singapore, Singapore; and, from 2008 to 2010, with the University of Waterloo, Waterloo, ON, Canada. He is currently a Full Professor with the Department of Control Science and Engineering and the Coordinator of the Networked Sensing and Control Group with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China. His research interests include estimation and control over sensor networks, sensor and actuator networks, and coverage and optimization in sensor networks.

Dr. Chen currently serves as an Associate Editor for several international journals, including the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTION SYSTEMS, the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, and IEEE NETWORKS. He is a Guest Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Computer Communication* (Elsevier), *Wireless Communication and Mobile Computer* (Wiley), and the *Journal of Network and Computer Applications* (Elsevier). He has also served as a Cochair of the Ad hoc and Sensor Network Symposium of the IEEE Global Communications Conference in 2011, the General Symposium Cochair of the Association for Computing Machinery International Wireless Communications and Mobile Computing Conference in 2009 and 2010, the Medium Access Control Track Cochair of the International Wireless Internet Conference in 2010, the Publicity Cochair of the IEEE International Conference on Mobile Ad hoc and Sensor Systems (IEEE MASS) in 2011, the Publicity Cochair of the IEEE International Conference on Distributed Computing in Sensor Sytems in 2011, the Publicity Cochair of the IEEE International Conference on Distributed Computing Systems (IEEE ICDCS) in 2012, and a Technical Program Committee member of IEEE ICDCS 2010, IEEE MASS 2010, the IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks in 2011, the IEEE International Conference on Computer Communications in 2011 (IEEE INFOCOM), IEEE INFOCOM 2012, IEEE ICDCS 2012, among others.



Hongyang Zhao received the B.E. degree in automation from Shanghai Jiaotong University, Shanghai, China, in 2011. He is currently working toward the M.S. degree with the Department of Control Science and Engineering, State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China.

He is currently a member of the Networked Sensing and Control Group, State Key Laboratory of Industrial Control Technology, Zhejiang University. His research interests include delay-tolerant networks and mobile social networks.

Wei Gao (M'12) received the B.E. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2005 and the Ph.D. degree in computer science from Pennsylvania State University, University Park, PA, USA, in 2012.

He is currently an Assistant Professor with the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA. His research interests include wireless and mobile network systems, mobile social networks, cyber–physical systems, and pervasive and mobile computing.

Peng Cheng (M'10) received the B.E. degree in automation and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2004 and 2009, respectively.

He is currently an Associate Professor with the Department of Control Science and Engineering, Zhejiang University. His research interests include networked sensing and control, cyber–physical systems, and robust control.

Dr. Cheng served as the Publicity Cochair for the IEEE International Conference on Mobile Ad hoc and Sensor Systems in 2013.