

Research Article

An Efficient Forwarding Capability Evaluation Method for Opportunistic Offloading in Mobile Edge Computing

Qian Wang ¹, Zhipeng Gao ¹, Kun Niu ², Yang Yang ¹, and Xuesong Qiu ¹

¹Network Management Center, The State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 10086, China

²School of Software Engineering, Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing 10086, China

Correspondence should be addressed to Zhipeng Gao; gaozhipeng@bupt.edu.cn

Received 13 April 2018; Revised 18 August 2018; Accepted 4 September 2018; Published 25 September 2018

Guest Editor: Kok-Seng Wong

Copyright © 2018 Qian Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Opportunistic offloading can be utilized to offload computing tasks and traffic data in Mobile Edge Computing (MEC). To improve the ratio of successful data offloading and reduce unnecessary data redundancy in opportunistic forwarding process, some methods of evaluating a device's forwarding capability are proposed. However, most of these methods do not consider the temporal impact from device mobility and the efficiency influence from the capability computation process. To settle these problems, we proposed a Transient-cluster-based Capability Evaluation Method (TCEM) to evaluate a device's data forwarding capability. The TCEM can be divided into two steps. The first step aims to reduce computational complexity by evaluating a device's possibility of contacting the destination within a time constraint based on the transient cluster generated by our proposed Transient Cluster Detection Method (TCDM). The second step is to calculate a device's probability of directly and indirectly forwarding data to the destination. The probability as a metric of evaluating a device's forwarding capability can be used in different data forwarding strategies. Simulation results demonstrate that the TCEM-based data forwarding strategy outperforms other data forwarding strategies from the aspect of the proportion of the data delivery ratio to the data redundancy.

1. Introduction

Opportunistic offloading as an emerging communication paradigm can be used in MEC to offload computation or traffic data by leveraging the opportunistic network formed by mobile devices [1–3]. For example, in computation offloading an edge mobile device with limited computing resources can offload computing tasks through opportunistic communication to other nearby devices with idle computing resources, and in traffic offloading the edge devices that cache the repeatedly requested or popular contents transfer the contents to subscribers by opportunistic communication to reduce network traffic and devices' energy consumption. Opportunistic offloading is based on opportunistic connection between mobile devices causing random delay; thus it is suitable for non-real contents like e-mail, podcast, and weather forecast, which can tolerate some delay in the delivery.

To maximally offload computing tasks and traffic data within a time constraint, the ratio of successful data delivery over opportunistic data forwarding needs to be maximized. In opportunistic offloading, the offloading device could forward the data to opportunistic encountered devices via store-carry-forward mechanism to increase the ratio of successful delivery. Forwarding data to every opportunistic encountered device until the deadline of the data or the destination is reached maximizes the data delivery ratio, but this process produces large amounts of redundant data carried by devices who cannot contact the destination within valid time of the data. These redundant data consumes large amounts of devices' limited resources, such as storage capacity and battery life. Therefore, in this paper, we study how to reduce the redundant data in the opportunistic forwarding process while maximizing the ratio of successful data delivery, when a device sends its caching data with a time constraint to another

requested device. The key challenge is how to make strategy on replicating data to opportunistic encountered devices.

Recent literatures make data forwarding strategies based on comparison of the device's data forwarding capability. Devices carried data only replicate the data to devices with higher capability or the device with the highest capability. Although data redundancy is reduced through this method, there exist some problems on evaluating the device's data forwarding capability. For example, the evaluation methods [4, 5] are derived from communities formed by the aggregation of contact information. However, contact patterns are time-varying. By aggregating contact information into an aggregated contact graph, some important contact information, e.g., the burst behaviour, may be lost, weakening the accuracy of evaluating forwarding capability. Additionally some evaluation methods such as EER [6] need to get global contact information in advance. It is unrealistic in large-scale mobility scenarios.

In this paper, to optimize data delivery ratio and to reduce unnecessary redundancy, we propose a distributed method—Transient-cluster-based Capability Evaluation Method (TCEM). The first step of TCEM is to judge the possibility of a device contacting the destination within the time constraint of the data using our proposed Transient Cluster Detection Method (TCDM). The second step is calculating the device's probability of encountering destination within the time constraint of the data as evaluation metric. The contributions of this method can be summarized as follows:

- (i) Our proposed TCEM is distributed. Every device uses this method to evaluate its forwarding capability based on the transient contact information it obtains. Comparing with some centralized methods, which need to get global contact information in advance, our proposed method is more feasible under mobility circumstance of the MEC.
- (ii) The proposed cluster detection method TCDM is formed by time-varying contact patterns of devices. Compared with other two transient cluster detection methods DRAFT [7] and CCM [8], TCDM is simpler and more efficient.
- (iii) Compared with some methods, which directly calculate every encountered device's probability as forwarding capability evaluation metric, our proposed TCEM effectively reduces the computational complexity because we first evaluate the encountered device's possibility of forwarding data to the destination, and only the device with high possibility calculates the probability.
- (iv) TCEM accurately evaluates a device's capability of forwarding data to a specific device within a time constraint, since it is based on individual pairs' Inter Contact Time (ICT) distributions and considers the influence of transient cluster's duration time in the probability calculation process.

The rest of the paper is organized as follows. Section 2 gives a brief overview of the related work. In Section 3, we give an overview of a TCEM-based data forwarding strategy.

Section 4 presents our proposed TCEM in detail. It contains TCDM and probability calculation. Section 5 evaluates the performance of our approach by simulation, and Section 6 concludes the paper.

2. Related Work

Opportunistic data forwarding protocol originates from Epidemic routing [9] which floods the network. Although this flooding-based algorithm can achieve the highest packet delivery ratio, it cause large amounts of redundant data copies in the network. Later studies devote to develop forwarding protocols to approach the performance of Epidemic routing with lower cost, which is measured by the number of data copies in the network. Currently, the most successful approaches for opportunistic forwarding are social-aware strategies [10]. The community structure formed by exploiting social contact patterns of nodes has been widely utilized in the methods of evaluating a node's forwarding capability, since it is more reliable and less susceptible to the randomness of human mobility.

In Bubble Rap [4], SimBet [11], MDDPC-based [12], and RPC-based [13] data forwarding strategies, the data forwarding capability of a node is represented by its social importance that is the degree of facilitating the communication among other nodes. They are evaluated based on communities. For increasing the efficiency of the decision making process of routing, some schemes combined features are proposed. For example, SimbetTS [14] adds the tie strength for utility calculation, Oi [15] and SCORP [16] combine users' social ties with their interest for social-aware opportunistic routing, and GROUPS-Net [17] combines social awareness with a probabilistic approach using group meetings as a measure of social context to improve data delivery ratio. SAMPLER [18] uses social communities and social popularity metrics as they were introduced in the original Bubble Rap [4] scheme and adds the individual mobility and points of interest within a region to them, helping to achieve high delivery ratio and reducing network overhead. LASS [19] is proposed taking into account the difference of members' internal activity within each community. It utilizes different levels of local activity within communities to realize efficient data forwarding. However, in these strategies the community is formed based on the previous cumulative contact knowledge. The transient characteristics of node contact pattern, which would influence data forwarding performance, are ignored.

References [6, 7, 20–24] consider time-varying contact pattern between nodes. Wei et al. [20] develop a novel evaluation method to analytically predict the forwarding capability of a mobile node based on a proposed transient community structure. An efficient temporal closeness and centrality-based data forwarding strategy TCCB [23] are proposed by predicting nodes' future temporal social contact patterns. However, these two schemes are unsuitable for our proposed situation. Since evaluating a node's data forwarding capability [20] is based on the centrality within its transient community and [23] is based on the centrality within all nodes in the network, they do not reflect the node's capability of forwarding data to a destination. The method DRAFT [7] measures the

forwarding capability of a node through judging the node's 2-hop cluster whether it contains destination or not. The cluster is formed based on participants' cumulative or decayed contact duration. SimBetAge [24] improves upon SimBet by adopting an aged graph to calculate the social metrics dynamically. These approaches are not efficient enough since they improve the performance of data forwarding by replicating data to all high possibility nodes, which still produce a lot of redundant data. TSM [21] evaluates a node's forwarding capability based on three types of time-varying social metrics: betweenness centrality, similarity, and tie strength, which are derived from analysing two sets of social data. Reference [22] proposes a CAOF scheme, which includes intercommunity and intracommunity phase. In the intercommunity phase, the node with higher global activeness and source-to-destination probability is selected to serve as the relay. Besides, in the intracommunity phase, the forwarding decisions are determined by the local metrics. EER [6] proposes to evaluate a node's forwarding capability based on its encounter value (EV), which is the number of directly encountering participants within the data's valid time. These four methods are unpractical in some cases, since they need to get global contact information within data's valid time in advance. Comparatively our proposed TCEM is distributed and effectively reduces the computational complexity because we first evaluate the encountered device's possibility of forwarding data to the destination and only the device with high possibility calculates the probability.

The second step of our proposed TCEM needs to calculate the probability that a node transfers data to a specific node within data's valid time. ICT is defined as the time elapsing between two consecutive encounters of two devices; thus it can be used to calculate the probability. Most probability calculation methods use the aggregate ICT distribution: power law + exponential tail or exponential, which is obtained by considering the samples from all pairs together. However Hernández-Orallo et al. [25] propose using the distribution of aggregate ICT to represent individual pairs' ICTs distributions will not be correct in general, if the network is heterogeneous. Hence in this paper we consider the difference of individual pairs' ICTs distributions in the probability calculation process for more accurately evaluating the forwarding capability of a node.

3. TCEM-Based Data Forwarding Strategy

In the process of opportunistic data forwarding, the device with data makes data forwarding strategy based on encountered devices' data forwarding capability. In fact the method of evaluating a device's data forwarding capability within time constraint is independent from data forwarding strategies. Different forwarding capability evaluation methods can be applied in the same data forwarding strategy. For example, [8, 20] all forward data to devices with higher capability, whereas they use different forwarding capability evaluation methods. In this section we apply our proposed device's data forwarding capability evaluation method TCEM in a data forwarding strategy that the carried data participant forwards data to participants with higher capability.

The data forwarding strategy is composed of two parts. First it uses the TCEM to evaluate the connected device's capability of forwarding data to the destination. Then it decides whether to replicate data to this device on the evaluation result. The detailed process is illustrated as follows.

When a device u_1 with data connects a device u_2 , u_1 first judges whether u_2 is carrying this data. If u_2 carries this data, u_1 directly skips u_2 . Otherwise u_1 executes the data forwarding strategy as illustrated in Figure 1.

In part A, the strategy uses TCEM to evaluate u_2 's data forwarding capability. The TCEM consists of two steps. In the first step, it judges whether u_2 has possibility of transferring data to the destination within the time constraint based on transient clusters. Only when u_2 's transient cluster or his adjacent transient clusters contain destination, the second step will be continued, otherwise u_1 will not replicate data to u_2 and the strategy will end. The detailed Transient Cluster Detection Method (TCDM) is demonstrated in Section 4.1. The second step is to use ICT distribution between participants to calculate the probability P_{u_2} that u_2 forwards data to destination within the time constraint of the data. The calculation process is shown in Section 4.2. In our proposed method the probability is used as the metric of evaluating a participant's forwarding capability.

In part B, the strategy determines whether replicating data to u_2 according to the forwarding probability of u_2 . If u_2 's probability of contacting destination within data's valid time is higher than u_1 , u_1 will replicate data to u_2 , or u_1 will not. Devices with data execute the afore-mentioned data forwarding strategy, when they opportunistically connect devices without carrying this data. The process is continued until the destination receives the data within data's valid time.

4. Data Forwarding Capability Evaluation Method of the Device (TCEM)

Our proposed TCEM consists of two steps: possibility evaluation and probability calculation. These two steps are based on transient clusters. In this section, we first describe our proposed Transient Cluster Detection Method TCDM, and how to evaluate a device's possibility of contacting destination (Section 4.1). We then elaborate on how to calculate the device's probability of successful data delivery within time constraint t (Section 4.2).

4.1. Transient Cluster Detection Method. In this paper, every device $u_i(1, 2, \dots, n)$ has a time-varying transient cluster T_i , which reflects a set of devices that have high contact rate with it, and the device u_i 's adjacent cluster A_i refers to transient clusters sharing the same device with u_i . As illustrated in Figure 2, a dashed circle represents a transient cluster. T_1 is the device u_1 's transient cluster at the time t . It contains u_2, u_3, u_4 which have high contact rate with u_1 at this moment, and T_2, T_3, T_4 are u_1 's adjacent clusters. We set a device that has a possibility to transfer data to a destination, only when its transient cluster or its adjacent transient clusters contain the destination. For example, in Figure 2 if the destination of the data is u_3 or u_5 , which is in u_1 's transient cluster or in u_1 's adjacent cluster T_2 , respectively, the device u_1 will

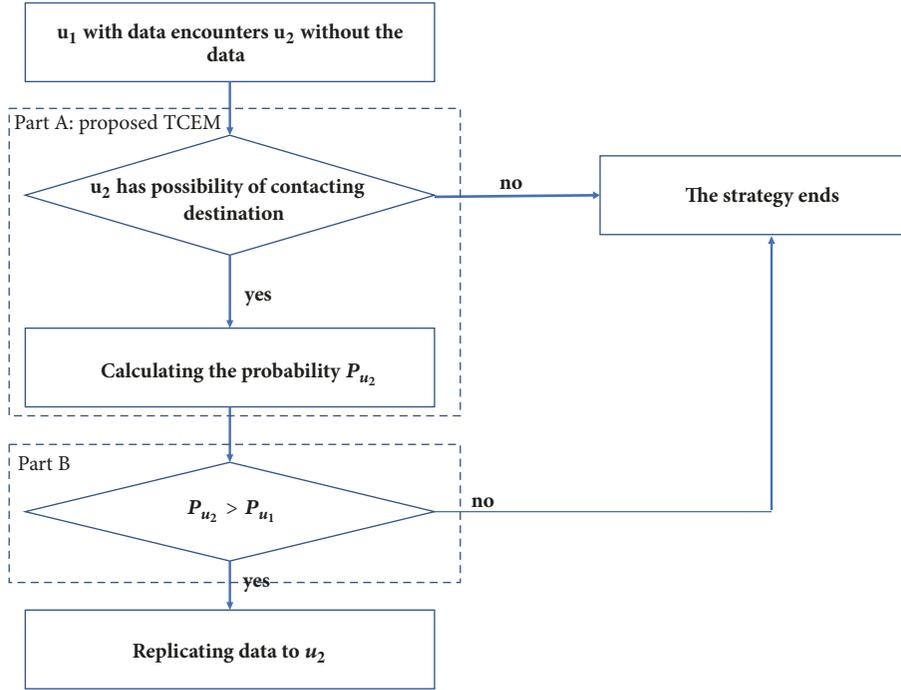


FIGURE 1: A TCEM-based opportunistic data forwarding strategy.

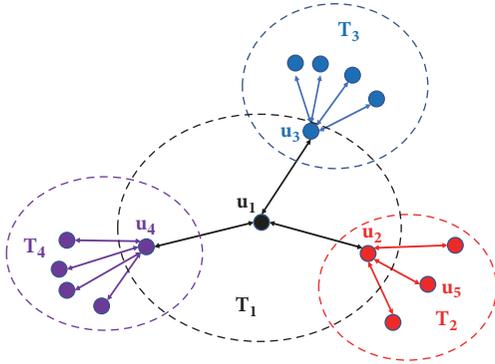


FIGURE 2: A device's transient cluster and its adjacent clusters.

have the possibility to transfer data to the destination. The evaluation standard is set because of two reasons. First the efficient content dissemination is mostly due to high contact rate nodes [26]. Second setting a device's neighbourhood view more than 2 hops does not improve the forwarding efficiency and even dramatically increases data redundancy [27].

Therefore, for evaluating a device u_i 's possibility of contacting a destination, the device also stores its adjacent clusters A_i , besides storing its transient cluster T_i . The contact rate used to construct transient clusters can be represented by ICT which is the time interval between two consecutive encounters. We define the contact rate between two devices is high if and only if the ICT between them is shorter than a predefined threshold x . A device would delete devices whose ICTs are bigger than x from its transient cluster. The value of x is set based on different traces. Since pairwise ICTs

are time-varying [21, 28], to accurately acquire the transient cluster and adjacent clusters of the current time, every x hours a device updates its transient cluster which only contains devices it encounters during this time period and updates its adjacent clusters when the device encounters a device. The process of building storage information of a device is described in Algorithm 1. It does not need centralized control and is independently performed by every device.

4.2. Calculating the Probability. After the first step of TCEM, a device's possibility of transferring data to destination has been determined. However, there is no guarantee that the data can be delivered to destination within a time constraint by the device, even though the path is at most 2 hops. To minimize the data redundancy, then in the second step, we calculate the device's probability of transferring data to destination within the time constraint, which represents the device's forwarding capability. The device with data only replicates data to encountered devices whose probabilities are higher than him. The calculation process is divided into two situations: (1) destination is in u_i 's transient cluster, and (2) destination is in u_i 's adjacent transient clusters. Table 1 gives the notations that are used in this section.

(1) The Destination in u_i 's Transient Clusters. We have defined that a device has possibility to transfer data to another device only if the path between them is no more than 2 hops. Hence u_i can transmit data to the destination d via two routes, as illustrated in Figure 3. The first route is that u_i directly transmits data to d . The second route is that u_i indirectly transmits data to d through a device u_j , who is a member of u_i 's transient cluster.

```

Input: Start time  $t_0$ .
Output:  $u_i$ 's transient cluster  $T_i$ ,  $u_i$ 's adjacent clusters  $A_i$ .
Initialize:  $T_i = \{\emptyset\}$ ,  $A_i = \{\emptyset\}$ ,  $t_j = 0$ .
For every encountered participant  $u_j$ ,  $u_i$  do
     $t_{ij} = t_{current} - t_j$  //  $t_{ij}$  is the ICT at the current time
    If ( $t_{ij} < x$ ) //  $u_j$  has high contact rate with  $u_i$ 
         $T_i = T_i + u_j$  // Adding encountered  $u_j$  to  $T_i$ 
         $t_j = t_{current}$  // Recording the encountered time
        If ( $A_i$  contains  $T_j$ ) // Whether encountered  $u_j$  before
            Update  $T_j$  in  $A_i$ 
        Else
             $A_i = A_i + T_j$ 
        Else
             $t_j = t_{current}$  // Recording the time that  $u_i$  encounters  $u_j$ 
    End
For ( $(t_{current} - t_0) \% x == 0$ ) do // Updating  $T_i$  every  $x$  hours
    For every  $u_j$  in  $T_i$  do
        If ( $(t_{current} - t_j) > x$ ) // The ICT is larger than  $x$  hour
             $T_i = T_i - u_j$  // Delete  $u_j$  from  $T_i$ 
             $A_i = A_i - T_j$  // Delete  $u_j$ 's transient cluster from  $A_i$ 
        End
    End
End

```

ALGORITHM 1: Building information stored by the device u_i .

TABLE 1: Symbols and their definitions.

The symbols	The definitions
$P_{u_i d}(t)$	the probability that u_i contacts d within t
$P_{u_i d}^d(t)$	the probability that u_i directly contacts d within t
$P_{u_i d}^{u_j}(t)$	the probability that u_i through a relay device u_j contacts d within t
$P_c(t_1, t_2)$	The probability of a transient cluster existing duration $t \in (t_1, t_2)$
$P_{u_i d}(t \leq t_v)$	u_i 's probability of contacting d before t_v
$P_{u_i d}^1(t_e, t_v)$	The probability of u_i contacting d when ($t_{ct} > t_v$)
$P_{u_i d}^2(t_e, t_v)$	The probability of u_i contacting d when ($t_{ct} \leq t_v$)
$P_{u_i d}'(t)$	the probability that u_i only indirectly contacts d within t
$P_{u_i d}'(t_e, t_v)$	The probability of u_i indirectly contacting d when ($t_{ct} > t_v$)
$P_{u_i d}'(t_e, t_v)$	The probability of u_i indirectly contacting d when ($t_{ct} \leq t_v$)
$P_{u_i d}'(t \leq t_v)$	u_i 's probability of only indirectly contacting d before t_v

In this paper, data forwarding depends on opportunistic connections between devices. We assume the size of sensed data is small, and it can be completely transferred during one contact and the transmission time can be neglected. Hence the device's probability of transferring data to destination within the time constraint is equal to the probability of contacting destination within the time constraint. The ICT distribution between devices in the real-world mobility traces shows some probability distribution models. We use individual pairwise ICT distribution to calculate the probability that a device directly or indirectly contacts destination within the time constraint and use f_{ij} to represent the ICT distribution between devices u_i and u_j .

The probability $P_{u_i d}(t)$ that u_i contacts d within time constraint t contains two parts: the probability of directly and indirectly contacts d . $P_{u_i d}(t)$ is calculated by

$$P_{u_i d}(t) = P_{u_i d}^d(t) + \sum_{u_j \in D} P_{u_i d}^{u_j}(t) \quad (1)$$

where

$$P_{u_i d}^d(t) = \int_0^t f_{u_i d} dt \quad (2)$$

$$P_{u_i d}^{u_j}(t) = \int_0^t \left(\int_0^{t_0} f_{u_i u_j} dt \int_0^{t-t_0} f_{u_j d} dt \right) dt_0$$

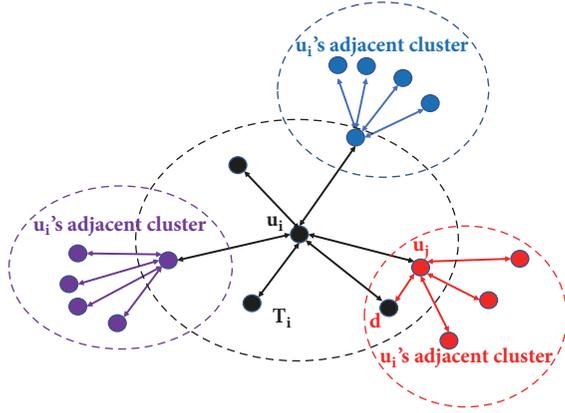


FIGURE 3: The destination is in a participant's transient cluster.

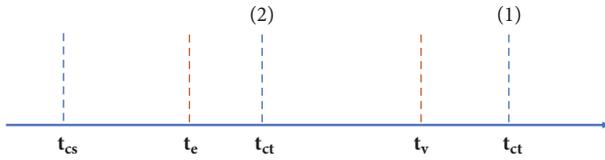


FIGURE 4: The relationship between cluster's duration and data's valid time.

$P_{u_i,d}^d(t)$ is the probability that u_i directly contacts d within t , $P_{u_i,d}^{u_j}(t)$ is the probability that u_i through a relay device u_j contacts d within t , and D is devices of u_i 's transient cluster whose transient cluster contains d .

Since our proposed calculation method is based on the transient cluster, the relationship between duration of the transient cluster and valid time of data would influence the value of the time constraint in the probability calculation process. As illustrated in Figure 4, t_{cs} is the start time of a transient cluster. t_{ct} is the end time of the transient cluster; thus $(t_{ct} - t_{cs})$ is the transient cluster's duration time. t_e is the time that the device carried data encounters another device. t_v is the deadline of the valid time of the data. The first relationship is (1) that the duration time of the cluster is longer than the valid time of the data. The second relationship is (2) that the duration time of the cluster is shorter than the valid time of the data, and under this situation the time constraint of the data becomes $(t_{ct} - t_e)$.

In the first relationship ($t_{ct} > t_v$), duration of the transient cluster is longer than valid time of the data. Equation (3) is the probability of this situation, where f_c is the distribution of the duration time of the transient cluster. In this situation, the real constraint time of the data is $(t_v - t_e)$. Equation (4) is the probability $P_{u_i,d}^1(t_e, t_v)$ that u_i contacts d within data's valid time under this situation. $P_{u_i,d}(t_v - t_e)$ is the probability that u_i contacts d within time constraint $(t_v - t_e)$.

To study f_c , we run the TCDM on two real datasets, Infocom6 [29] and Cambridge [30]. The detailed information about these two datasets is illustrated in Table 1. We observe the duration time of a transient cluster on daily basis and find its distribution can be approximated by exponential distribution. We take one transient cluster in Cambridge trace

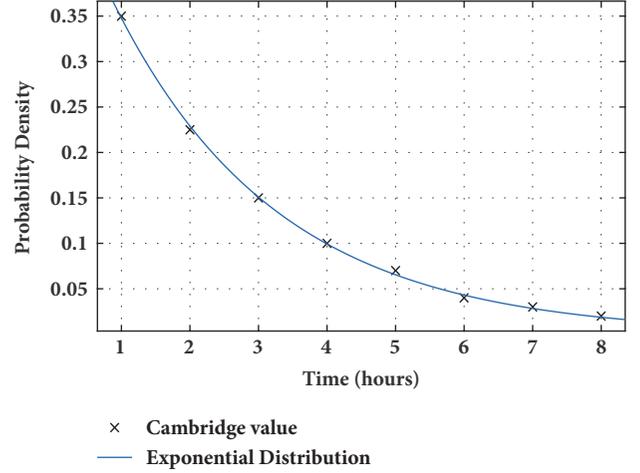


FIGURE 5: The PDF of a transient cluster's duration time.

as an example, shown in Figure 5. The approximation does not seem perfect, since the samples used to train the distribution are limited. It should be better if more data are used.

$$P_c(t_v - t_{cs}, \infty) = \int_{t_v - t_{cs}}^{\infty} f_c dt \quad (3)$$

$$P_{u_i,d}^1(t_e, t_v) = P_{u_i,d}(t_v - t_e) P_c(t_v - t_{cs}, \infty) \quad (4)$$

The second relationship ($t_{ct} \leq t_v$) is that duration of the transient cluster is smaller than valid time of the data. Equation (5) is the probability of this situation. In this situation, the real constraint time of the data is $(t_{ct} - t_e)$. Equation (6) is the probability $P_{u_i,d}^2(t_e, t_v)$ that u_i contacts d within data's valid time under this situation. $P_{u_i,d}(t_{ct} - t_e)$ is the probability that u_i contacts d within time constraint $(t_{ct} - t_e)$.

$$P_c(t_e - t_{cs}, t_v - t_{cs}) = \int_{t_e - t_{cs}}^{t_v - t_{cs}} f_c dt \quad (5)$$

$$P_{u_i,d}^2(t_e, t_v) = P_{u_i,d}(t_{ct} - t_e) P_c(t_e - t_{cs}, t_v - t_{cs}) \quad (6)$$

In summary, (7) is the probability $P_{u_i,d}(t \leq t_v)$ that u_i contacts the destination before the expiration of the data's valid time.

$$P_{u_i,d}(t \leq t_v) = P_{u_i,d}^1(t_e, t_v) + P_{u_i,d}^2(t_e, t_v) \quad (7)$$

(2) *The Destination in u_i 's Adjacent Transient Clusters.* We discuss the situation that the destination d is not in u_i 's transient cluster, but it is in u_i 's adjacent transient clusters. Figure 6 shows an example that d is in the u_i 's adjacent cluster T_j . u_i can transmit data to d through u_j . The probability $P_{u_i,d}^l(t)$ that u_i encounters destination d within time constraint t only by indirect contacting is calculated in (8), where T is devices of u_i 's transient cluster whose transient clusters contain d .

$$P_{u_i,d}^l(t) = \sum_{u_j \in T} P_{u_i,d}^{u_j}(t) \quad (8)$$

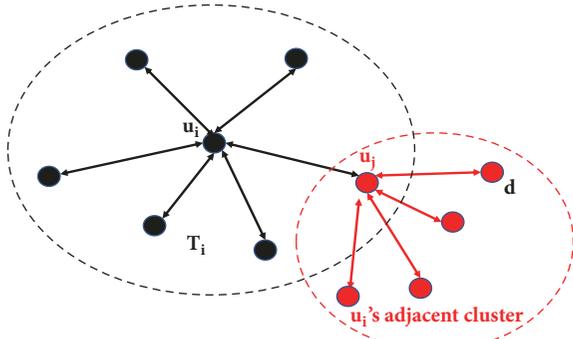


FIGURE 6: The destination is in a device's adjacent cluster.

The probability calculation is based on transient clusters. The relationship between the time, during which destination is in the opportunistic connected device's adjacent transient cluster, and the valid time of the data causes two possibilities. It is similar with Section 4.2(1).

When $(t_{ct} > t_v)$, (9) is the probability $P_{u_i,d}^1(t_e, t_v)$ that u_i encounters d , under this situation that the duration of transient cluster is longer than data's valid time, where $P_{u_i,d}^1(t_v - t_e)$ is calculated based on (8).

$$P_{u_i,d}^1(t_e, t_v) = P_{u_i,d}^1(t_v - t_e) \int_{t_v - t_{cs}}^{\infty} f_c dt \quad (9)$$

When $(t_{ct} \leq t_v)$, (10) is the probability $P_{u_i,d}^2(t_e, t_v)$ that u_i encounters destination d , under this situation that the duration of transient cluster is smaller than valid time of the data, where $P_{u_i,d}^1(t_{ct} - t_e)$ is calculated based on (8). In summary, (11) is the probability that u_i contacts destination d before the expiration of valid time of the data.

$$P_{u_i,d}^2(t_e, t_v) = P_{u_i,d}^1(t_{ct} - t_e) \int_{t_e - t_{cs}}^{t_v - t_{cs}} f_c dt \quad (10)$$

$$P_{u_i,d}^1(t \leq t_v) = P_{u_i,d}^1(t_e, t_v) + P_{u_i,d}^2(t_e, t_v) \quad (11)$$

5. Performance Evaluation

Our proposed TCEM is a method of measuring a device's data forwarding capability. Since it is based on the transient cluster, the performance of forming transient clusters will influence the performance of the method. Therefore we evaluate the performance of the Transient Cluster Detection Method (TCDM) of TCEM. Meanwhile, we evaluate the performance of a TCEM-based data forwarding strategy. In the experiments, the data delivery ratio is the proportion of data items, which are successfully delivered to the destination through opportunistic forwarding before data expires. It is improved by increasing the number of data copies, which represented the network overhead. Hence, to reflect the efficiency, the performance of the data forwarding strategy is evaluated by the proportion of data delivery ratio to the network overhead.

TABLE 2: Trace summary.

	Infocom6	Cambridge
Environment	Conference	City
Duration (days)	3	12
Device number	78	36
Inter-probe time/s	120	600

Our experiments are based on two real datasets, Infocom6 [29] and Cambridge [30]. These datasets are formed based on the devices that periodically detect their peers via Bluetooth interfaces, and a contact is recorded when two devices move into the communication range of each other. The details of two datasets are shown in Table 2.

5.1. Transient Cluster Evaluation. We compare the performance of TCDM with two transient cluster detection methods: Contact-burst-based Clustering Method (CCM) [8] and Distributed Rise and Fall spatio-Temporal (DRAFT) clustering method [7]. Besides the distributed feature and the feasibility of the method, the size of clusters detected by the method would also influence efficiency of the data forwarding strategy, as a data forwarding strategy decides whether to forward data to an encountered participant by searching all numbers of this participant's cluster. Hence in this section we compare the performance of methods from three metrics: mean cluster size, max cluster size, and the used cluster size of evaluating forwarding capability.

CCM. It detects transient clusters by clustering pairs of nodes with similar contact bursts together. A contact burst between two nodes refers to a period when contacts frequently appear between these two nodes. The transient cluster detected by CCM is time-varying, and the cluster's forming process needs central control.

DRAFT. It is distributed. Every node forms a cluster based on cumulated or decayed contact duration. It uses three parameters τ , t , δ to govern the rate at which clusters grow and decay. If the cumulated contacted duration is longer than τ , the encountered device will be added to the cluster. Otherwise the device will be deleted from the cluster. A time frame of length t seconds governs the interval at which the cumulative durations for each device are decayed. At the end of each time frame, the contacted duration is decreased by multiplying the parameter δ .

Our proposed TCDM is distributed. Every device forms a time-varying transient cluster, which contains a set of devices that have high contact rate with him. In this paper, we set a device that will add an encountered device to his transient cluster, only if these two devices' intercontact time is lower than predetermined parameter x . The value of x is set empirically based on traces.

In these two datasets, we found that each pairwise contact is a series of contact bursts during which two devices' contact rate is high, and two devices' intercontact time is lower than 1 hour in these contact burst periods. For example, we choose any two devices from the Cambridge dataset. The contact is

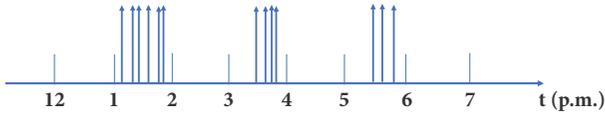


FIGURE 7: Contact rate of two nodes.

illustrated in Figure 7. An arrow represents a contact between two devices. Thus we set $x = 1$. If these two devices do not contact within one hour that means the two devices' contact rate is not high, and a device will be deleted from another device's transient cluster.

Mean Cluster Size. Figure 8 shows that the mean cluster size of TCDM is the smallest among these three methods. CCM's mean cluster size is the biggest, and it is almost three times bigger than TCDM. We illustrate the reasons as follows.

In the DRAFT method, a cluster is a device-centric one-hop cluster, but a device is added to or deleted from a cluster based on the cumulative or decayed encounter duration time. In these two datasets, the cluster of DRAFT will contain high contact rate devices. For example, in Infocom6 dataset, two devices with high contact rate mean these two devices in a contact burst period. Mean contact duration between devices is 24 s. Two participants contact 6 times during a contact burst period. Therefore, in a contact burst period, two devices' cumulative encounter duration time is 144 s, which exceeds the predetermined threshold 120 s. The device will be added to another device's cluster. However, since a device is deleted from a cluster based on the decayed duration time, it might lead to a cluster that still accumulated some devices which do not have high contact rate with the central device now, whereas members of a TCDM cluster are devices, which have high contact rate with a central device. Therefore, the cluster detected by DRAFT is bigger than TCDM.

The mean size of DRAFT's cluster is smaller than CCM, because a cluster detected by DRAFT method aims to detect devices which have high contact frequency with a central device during a period. It is a one-hop cluster. However in CCM any two devices have high contact rate duration; this period will be added to a cluster. The hops do not have limitation. It is absolutely bigger than a DRAFT cluster.

In summary, comparing with DRAFT, although the complexity of TCDM's cluster forming method is similar, TCDM's cluster size is smaller than DRAFT. Comparing with CCM, besides the cluster size TCDM is more feasible and efficient. We demonstrate it from two aspects. First CCM needs to get all devices' contact information in advance. Second in the process of forming cluster, CCM has to traverse all clusters to merge similar clusters. This process wastes large amounts of computing resources and time. However, our proposed TCDM is distributed. Every device forms his transient cluster only based on two devices' real-time contact frequency. Therefore, comparing with CCM, TCDM is more feasible and efficient.

Max Cluster Size. Figure 9 shows the max cluster detected by these three methods. The max cluster detected by TCDM

and DRAFT method is smaller than cluster detected by CCM. CCM's max cluster contains more than half of the whole devices. It would extremely increase computation complexity.

The Cluster Size of Evaluating Forwarding Capability. DRAFT-based and our proposed TFCM-based data forwarding strategy determine whether forwarding data to an encountered device is based on its 2-hop cluster. Therefore, a device's 2-hop cluster size is very important for the performance of these two methods. However, CCM-based data forwarding strategy is based on a cluster. Hence, we compare mean 2-hop cluster size of TCDM and DRAFT, and mean cluster size of CCM. Figure 10 shows our proposed TCDM's cluster is the smallest among these three cluster detection methods. It includes more or less 30% of all devices, but DRAFT's 2-hop cluster contains almost half of all devices and CCM's cluster contains more than half of all devices. The reasons are illustrated as follows. First, since the 1-hop cluster of TCDM is smaller than 1-hop cluster of DRAFT, 2-hop cluster of TCDM is smaller than 2-hop cluster of DRAFT. Second, in DRAFT, a cluster contains 2-hop participants but a cluster of CCM does not limit hops; hence the cluster size of CCM is bigger than DRAFT.

In conclusion, first, the cluster size of our proposed TCDM is the smallest. It makes the efficiency of judging the forwarding capability of a device by searching all members in the cluster the best. In addition, comparing with CCM, our proposed TCDM is more feasible and more efficient because of its distribution and simplicity of cluster formation method. Comparing with DRAFT's cluster, which cannot accurately express the change of contact rate between participants, TCDM can reflect the change of devices' contact in real time. Since evaluating a participant's data forwarding capability is based on contact rate between devices, TCDM is more accurate in the evaluation of a device's forwarding capability.

5.2. Data Forwarding Performance Evaluation. In the experiment for fairness sources and destinations are picked randomly, and the data's generation time is randomly chosen in the daytime because nodes' activity time is low at night, which may result in inaccurate comparison. The performance is measured by the metric: the proportion of data delivery ratio to network overhead. Our proposed TCEM-based data forwarding strategy compares with Epidemic [9], Bubble Rap [4], DRAFT-based [7], and CCM-based [8] data forwarding strategy.

Epidemic. The data item is forwarded to every encountered device without the same data. It serves as upper bound.

Bubble Rap. This strategy utilizes both centrality and community. CPM (K-clique) is used to detect communities. The data item is always forwarded to a higher centrality device, until it reaches a device that belongs to the same community as the destination. When the data item reaches the destination community, it is forwarded to higher centrality device within the community's scope until the destination is reached.

DRAFT-Based Data Forwarding Strategy. The data will be replicated to encountered devices whose 2-hop clusters contain the destination until a carried data device encounters destination.

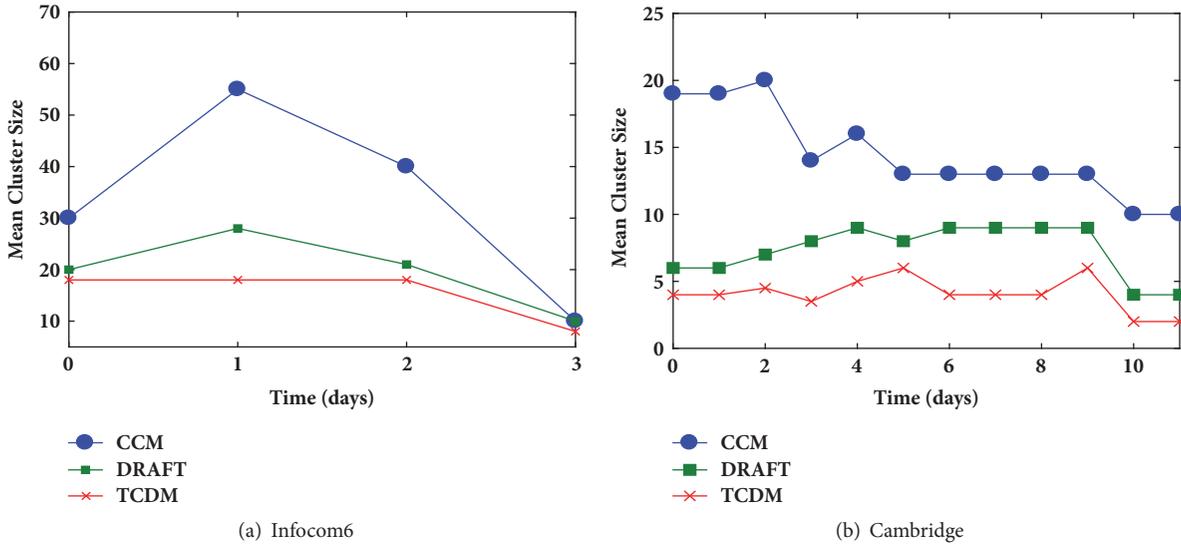


FIGURE 8: Mean cluster size.

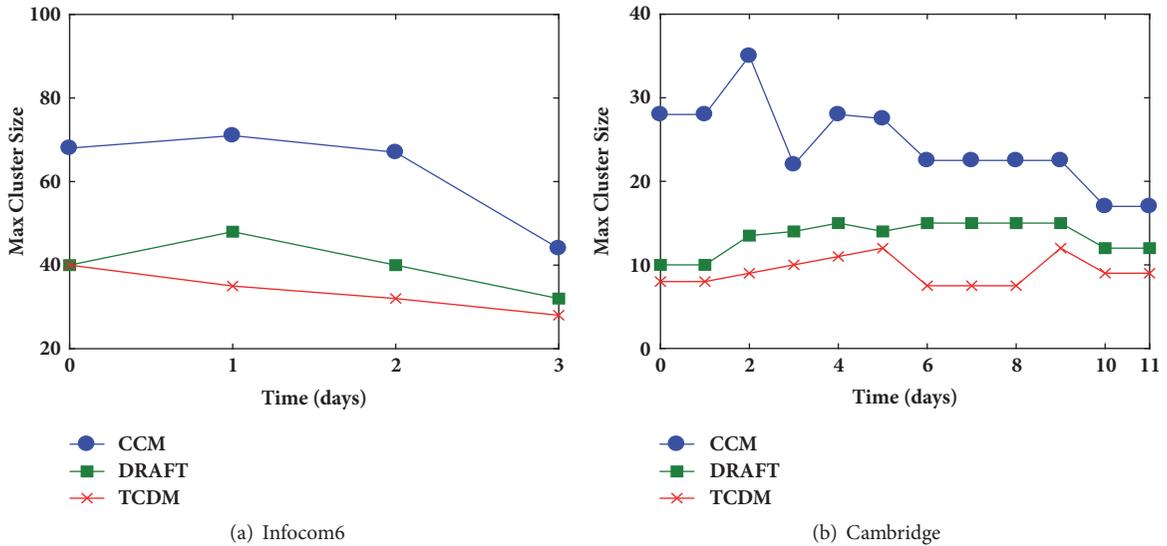


FIGURE 9: Max cluster size.

CCM-Based Data Forwarding Strategy. It utilizes transient cluster (TC) as the forwarding unit, and data is always forwarded to the TC with better relaying capability to the destination within data's time constraint. Once data reaches a new TC with larger relaying capability, the data is distributed to all nodes met in the TC. The relaying capability of the current TC is computed by summing the probability that each participant of the TC appears in destination's transient cluster within data's valid time. The carried data device deletes data when he has neither encountered a device in TC with larger relaying capability nor gone to a TC with larger relaying capability in an appointed time period. The above-mentioned process ended when data reached the destination or the time exceeds data's valid time.

The result is shown in Figure 11. Considering the delivery ratio and the overhead the performance of Bubble Rap is

the worst comparing with TCEM-based, DRAFT-based, and CCM-based data forwarding strategies. The reason is that these three strategies consider device's contact rate is time-varying, but Bubble Rap uses aggregate contact information which cannot reflect device's contact rate within data's valid time. Comparing with other three data forwarding strategies, Bubble Rap cannot accurately evaluate device's forwarding capability within the time constraint. Thus, its data forwarding performance is the worst.

Comparing with DRAFT-based and CCM-based data forwarding strategies, our proposed TCEM-based data forwarding strategy's delivery ratio is the lowest, but at the same time its overhead is also the lowest as illustrated in Figures 12 and 13. Integrating these two parameters its proportion of data delivery ratio to network overhead is highest. First, data forwarding of DRAFT and TCEM is based on the

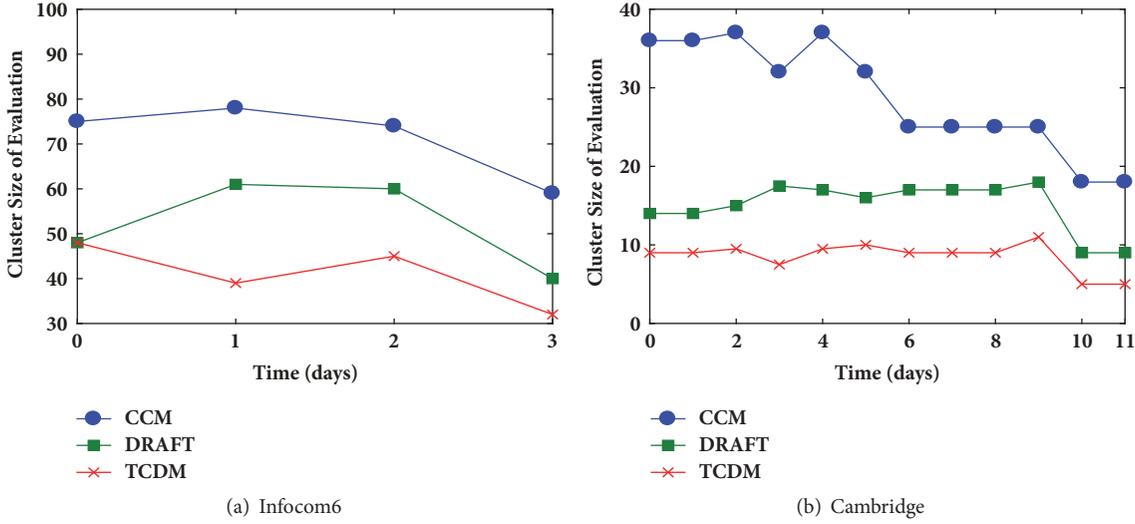


FIGURE 10: The cluster size of evaluating forwarding capability.

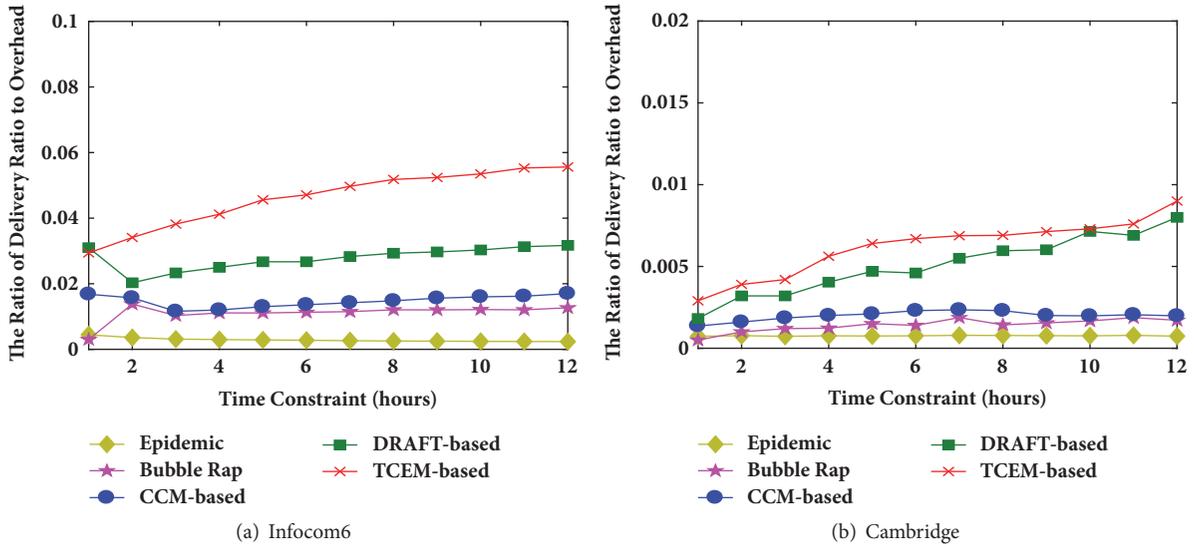


FIGURE 11: The proportion of delivery ratio to overhead.

encountered device’s 2-hop transient cluster. DRAFT’s 2-hop cluster size is larger than TCEM’s 2-hop cluster, and it contains all members of TCEM’s cluster. In the process of data forwarding, DRAFT forwards data to all encountered devices whose 2-hop cluster contains destination, but TCEM forwards data to an encountered device whose 2-hop cluster contains destination and probability of encountering destination is higher than the carried data participant. TCEM-based data forwarding strategy reduces some unnecessary overhead of replicating data to devices whose probability of encountering destination is low. Thus, although TCEM-based data forwarding strategy’s data delivery ratio is still a little lower than DRAFT-based data forwarding strategy, it greatly reduces overhead.

Second, data delivery ratio and overhead of CCM-based data forwarding strategy are the highest. Comparing with

DRAFT and TCEM only forward data to the encountered device, CCM’s data forwarding is based on the transient cluster. The data is distributed to all devices encountered in the transient cluster. Meanwhile, CCM’s transient cluster is the largest. It contains all devices that directly or indirectly and frequently contact with the encountered device. Considering these two reasons, CCM’s data delivery ratio is the highest, but its overhead is also high. At the same time, since in CCM strategy the carried data device deletes data when he has neither encountered a device in TC with larger relaying capability nor gone to a TC with larger relaying capability in every appointed time period, it will lead to the increase of the overhead as data may be copied many times to the same device. Hence, CCM’s overhead is extremely high and its proportion of data delivery ratio to network overhead is the lowest. Moreover, the computation complexity of

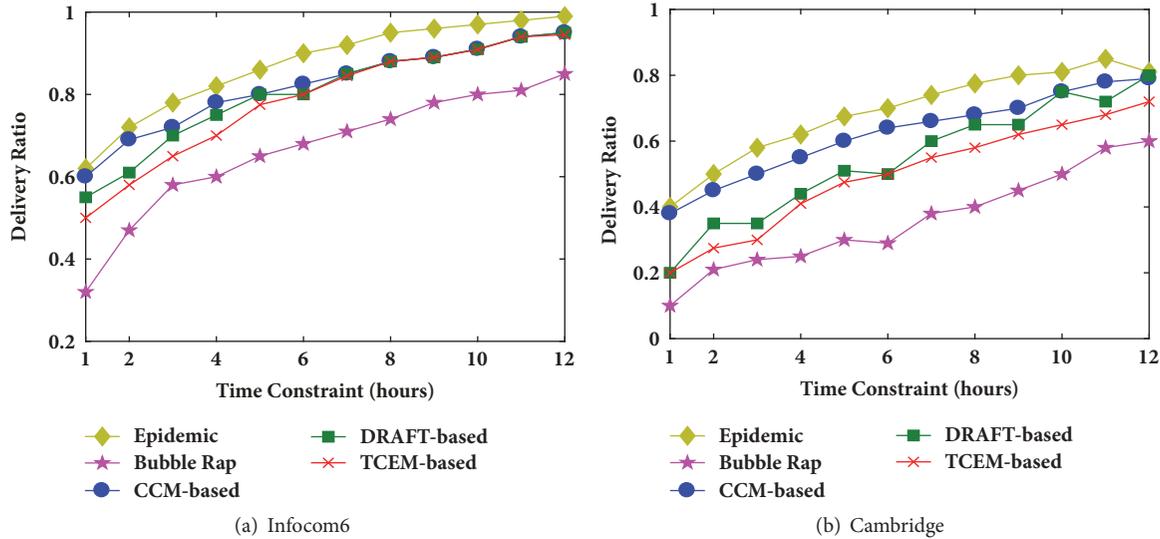


FIGURE 12: Data delivery ratio.

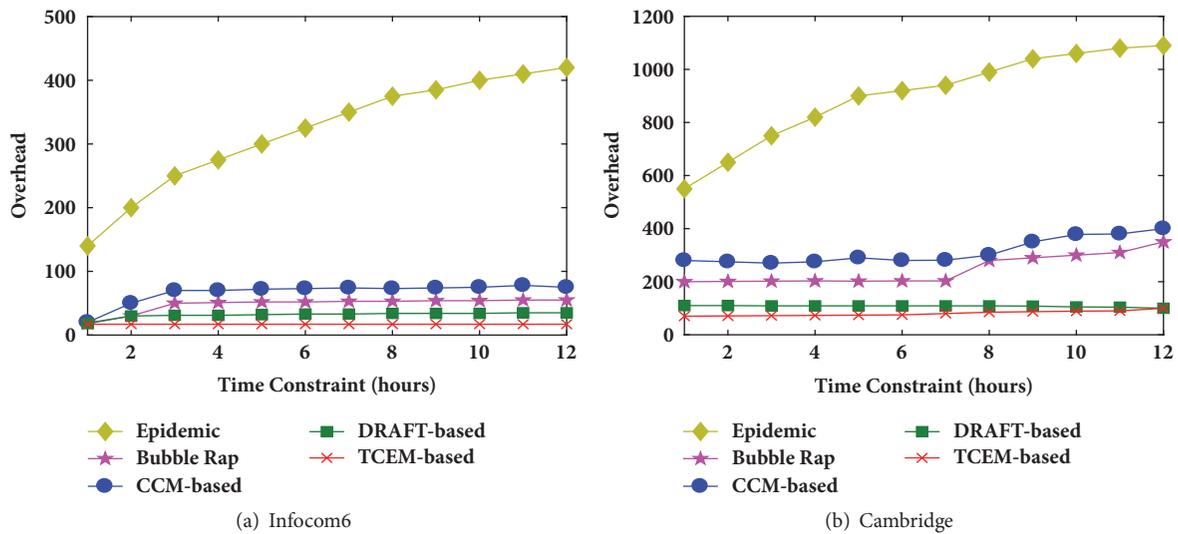


FIGURE 13: The network overhead.

CCM-based data forwarding strategy is greatly high, because it needs to find every encountered device’s current TC and calculate the probability that every member of TC encounters destination.

In conclusion, integrating delivery ratio and overhead the data forwarding performance of Bubble Rap is the worst. Comparing with DRAFT-based and CCM-based data forwarding strategies, although our proposed TCEM-based data forwarding strategy’s data delivery ratio is the lowest, taking account of delivery ratio and network overhead, the performance of our proposed method is the best.

6. Conclusions

In this paper, we proposed a method TCEM to evaluate the device’s data forwarding capability. The method is based

on the time-varying transient cluster and we also propose a method TCDM to detect the transient cluster. Simulation results show that the data forwarding strategy based on our proposed TCEM outperforms other existing data forwarding approaches.

Data Availability

The Infocom6 and Cambridge datasets supporting this meta-analysis are from previously reported studies and datasets, which have been cited.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by National Science & Technology Pillar Program (2015BAH03F02) and National Key Research and Development Program of China (2016YFE0204500).

References

- [1] D. Xu, Y. Li, X. Chen et al., "A survey of opportunistic offloading," *IEEE Communications Surveys & Tutorials*, 2018.
- [2] Y. Mao, C. You, J. Zhang et al., "A survey on mobile edge computing: The communication perspective," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.
- [3] Y. Han, T. Luo, D. Li, and H. Wu, "Competition-based participant recruitment for delay-sensitive crowdsourcing applications in D2D networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 12, pp. 2987–2999, 2016.
- [4] P. Hui, J. Crowcroft, and E. Yoneki, "BUBBLE Rap: social-based forwarding in delay-tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 11, pp. 1576–1589, 2011.
- [5] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation forwarding," in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '08)*, pp. 251–260, May 2008.
- [6] H. Chen and W. Lou, "Contact expectation based routing for delay tolerant networks," *Ad Hoc Networks*, vol. 36, pp. 244–257, 2016.
- [7] M. Orlinski and N. Filer, "The rise and fall of spatio-temporal clusters in mobile ad hoc networks," *Ad Hoc Networks*, vol. 11, no. 5, pp. 1641–1654, 2013.
- [8] X. Zhang and G. Cao, "Transient community detection and its application to data forwarding in delay tolerant networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 2829–2843, 2017.
- [9] W. Mitchener and A. Vadhat, "Epidemic routing for partially connected ad hoc networks," Technical Report CS-2000-06, 2000.
- [10] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware D2D communications: qualitative insights and quantitative analysis," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 150–158, 2014.
- [11] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant MANETs," in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'07)*, pp. 32–40, September 2007.
- [12] H. Zhou, L. Tong, T. Jiang, S. Xu, J. Fan, and K. Lv, "Maximum data delivery probability-oriented routing protocol in opportunistic mobile networks," *Peer-to-Peer Networking and Applications*, vol. 10, no. 3, pp. 500–509, 2017.
- [13] J. Wu, J. Wang, L. Liu, M. Tanha, and J. Pan, "A data forwarding scheme with reachable probability centrality in DTNs," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '16)*, pp. 1–6, April 2016.
- [14] E. M. Daly and M. Haahr, "Social network analysis for information flow in disconnected delay-tolerant MANETs," *IEEE Transactions on Mobile Computing*, vol. 8, no. 5, pp. 606–621, 2009.
- [15] L. Amaral, R. Sofia, P. Mendes, and W. Moreira, "Oi! - Opportunistic data transmission based on Wi-Fi direct," in *Proceedings of the IEEE INFOCOM - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS '16)*, pp. 578–579, San Francisco, Calif, USA, April 2016.
- [16] W. Moreira, P. Mendes, and S. Sargento, "Social-aware opportunistic routing protocol based on users interactions and interests," in *Proceedings of the International Conference on Ad Hoc Networks*, pp. 100–115, Springer, 2013.
- [17] I. O. Nunes, C. Celes, P. O. S. Vaz de Melo, and A. A. F. Loureiro, "GROUPS-NET: Group meetings aware routing in multi-hop D2D networks," *Computer Networks*, vol. 127, pp. 94–108, 2017.
- [18] I. O. Nunes, C. Celes, I. Nunes, P. O. S. Vaz De Melo, and A. A. F. Loureiro, "Combining spatial and social awareness in D2D opportunistic routing," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 128–135, 2018.
- [19] Z. Li, C. Wang, S. Yang, C. Jiang, and X. Li, "LASS: Local-activity and social-similarity based data forwarding in mobile social networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 1, pp. 174–184, 2015.
- [20] W. Gao, G. Cao, T. La Porta, and J. Han, "On exploiting transient social contact patterns for data forwarding in delay-tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 1, pp. 151–165, 2013.
- [21] D. Xie, X. Wang, L. Liu, and L. Ma, "Exploiting time-varying graphs for data forwarding in mobile social Delay-Tolerant Networks," in *Proceedings of the 24th IEEE/ACM International Symposium on Quality of Service (IWQoS '16)*, June 2016.
- [22] J. Tao, H. Wu, S. Shi, J. Hu, and Y. Gao, "Contacts-aware opportunistic forwarding in mobile social networks: A community perspective," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '18)*, pp. 1–6, April 2018.
- [23] H. Zhou, V. C. M. Leung, C. Zhu, S. Xu, and J. Fan, "Predicting temporal social contact patterns for data forwarding in opportunistic mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10372–10383, 2017.
- [24] J. Á. B. Link, N. Viol, A. Goliath, and K. Wehrle, "SimBetAge: utilizing temporal changes in social networks for pocket switched networks," in *Proceedings of the 1st ACM Workshop on User-Provided Networking: Challenges and Opportunities (U-NET '09)*, pp. 13–18, ACM, New York, NY, USA, December 2009.
- [25] E. Hernández-Orallo, J. C. Cano, C. T. Calafate, and P. Manzoni, "New approaches for characterizing inter-contact times in opportunistic networks," *Ad Hoc Networks*, vol. 52, pp. 160–172, 2016.
- [26] A.-K. Pietiläinen and C. Diot, "Dissemination in opportunistic social networks: the role of temporal communities," in *Proceedings of the 13th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '12)*, pp. 165–174, June 2012.
- [27] T. Phe-Neau, M. Dias De Amorim, and V. Conan, "The strength of vicinity annexation in opportunistic networking," in *Proceedings of the 32nd IEEE Conference on Computer Communications (IEEE INFOCOM '13)*, pp. 3369–3374, April 2013.
- [28] S. C. Nelson, M. Bakht, R. Kravets, and A. F. Harris, "Encounter-based routing in dtns," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 13, no. 1, pp. 56–59, 2009.
- [29] N. Eagle and A. Pentland, "Reality mining: sensing complex social systems," *Personal and Ubiquitous Computing*, vol. 10, no. 4, pp. 255–268, 2006.
- [30] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace - Hagggle," 2009, <http://crawdad.cs.dartmouth.edu/cambridge/hagggle>.



Hindawi

Submit your manuscripts at
www.hindawi.com

