Cluster Based VDTN Routing Algorithm with Multi-attribute Decision Making

Songjie Wei, Qianrong Luo(✉), Hao Cheng, and Erik Joseph Seidel

School of Computer Science and Engineering, Nanjing University of Science and Technology, 200 Xiaolingwei, Nanjing 210094, China
{swei,15205170353,hcheng,seidel}@njust.edu.cn

Abstract. Multi-copy routing may cause the excessive consumption of network resources in Vehicular Delay Tolerant Network (VDTN). Most existing routing algorithms only consider controlling the redundant messages in the global network to improve performance, but the local road redundancy still exists. In this paper, we propose a cluster based VDTN routing algorithm with multi-attribute decision-making (CR-MA) in urban areas. CR-MA enables nodes aware of local message distribution. We consider both the message distribution as coverage and the node and network attributes to establish a multi-attribute decision-making model which manages to achieve tradeoff between message redundancy and delivery performance. CR-MA is benchmarked in simulation against Prophet, Epidemic and SAW. Experiment results proves its superiority in both message routing effective, performance, and overhead.

Keywords: VANET · DTN routing · Multi-attribute decision making

1 Introduction

Vehicular Ad-hoc Network (VANET) is a typical application scenario of Delay Tolerant Network (DTN) [1] for inter-vehicle communication using wireless communication technology. It is difficult to ensure a stable communication path between source and destination due to the high mobility of nodes, intermittent inter-node connectivity and the high change of network topology. We study the Vehicular DTN (VDTN) that extends the features of DTN to VANET. The key issue in VDTN routing algorithm is how to choose a suitable relay node for each message based on the paradigm of store carry and forward (SCF).

Existing VDTN routing protocols can be classified as either single-copy or multiple-copy routing protocols based on the number of message copies disseminated through the network. Direct Delivery [2] and First Contact [3] are typical single-copy routing protocols. A single-copy routing protocol's network overhead is trivial but it suffers long delivery delay and low delivery ratio. A message should be replicated to enough nodes to increase the chance of being exposed to its destination node. Epidemic [4] is a multiple-copy routing protocol implementing flooding to minimize the delivery delay and maximize the delivery ratio when nodes having enough buffer space. Spray and Wait [5], Binary Spray and Wait [5] and Spray and Focus [6] are three typical controlled

Á. Rocha et al. (eds.), *Recent Advances in Information Systems and Technologies*,
Advances in Intelligent Systems and Computing 570, DOI 10.1007/978-3-319-56538-5_34

flooding routing protocols that set the maximum number of replications. Prophet [7] and MaxPro [8] are information-based routing protocols that improve routing performance by using history encounter times of nodes. The information used currently can be divided into three types: history encounter times of nodes, prior information of the network and location information. However, none of those considers both the message distribution and node's attributes to tune message replication.

Social-based routing protocols are an evolution for DTN by exploring social behaviors and properties of nodes to group them into different clusters or communities. Contact frequency and duration [9], movement area of the node [10] and the trajectory of node [11] are three bases for most clustering methods.

Currently, most routing algorithms only consider controlling the redundant messages in the global network to improve performance, but the local network congestion problem still exists. The heterogeneity of the message coverage on each road make the volume of message exchanged highly diversified upon node encounters at different geographic locations and time. Therefore, the local message coverage should be carefully considered as an impact factor in the routing algorithm in order to control local network congestion. Existing clustering methods mainly distribute nodes into two-dimensional clusters. We believe the structure of urban road limits nodes' direction and scope, and propose a one-dimensional linear clustering method based on road map.

In this paper, we propose CR-MA, a cluster-based multi-attribute algorithm for message forwarding in VDTN. First, one-dimensional linear clusters are formed based on the road structures. Then nodes try to understand and predict the message coverage in each cluster. We consider both the message distribution and node attributes when deciding when and where to duplicate messages.

2 System Model

The application scenario of a VDTN routing algorithm involves the configuration, wireless communication capability, and movement of each individual vehicle node, as well the organization of all the nodes as a communication network. This section introduces how these essential elements are modeled and presented in the proposed message distribution and routing algorithm. Vehicle nodes are grouped into clusters based on their locations and moving status. We present a road-based node clustering mechanism and explain how messages are transmitted intra and inter-clusters.

2.1 Vehicular Node

A vehicular node moving on road can initiate, buffer, transmit and deliver messages. A message is generated in fixed size by the initiator when needed by an application running on the node. Messages whose destinations are different from the current node will be buffered on the node and forwarded to the next intermediate hop when appropriate. There is a constraint on the message buffer size, and thus restrict the number of live messages (messages not at destination) buffered on each node. Neighbor nodes when in each other's wireless communication range may reliably exchange information about each other's moving status, message buffer digest, and forward messages when necessary. Each node makes routing and forwarding decisions independently by considering two factors, the nodes' intrinsic state, such as power and storage resources, moving history and status, and external message and network state, including message coverage and neighbor connections. A vehicular node also has awareness of its current road location, i.e. which road it is current running on, either by road id or name. While an onboard GPS plus electrical map is sufficient to provide such information, the road-awareness requirement is easier to be met with other vague methods like cellular base station sensing or road nameplate recognition.

The node movement model used in this paper is the Shortest Path Map Based movement model (SPMB). Initially, all vehicular nodes are uniformly distributed on the urban roads. Each node randomly selects a point in the road map as the current destination. The roads from current location along the shortest path to the destination location are calculated by the Dijkstra algorithm to make a route. Once the destination is reached, node randomly selects the next destination to repeat the route navigation.

2.2 Vehicular Network

All the vehicular nodes equipped with wireless communication and ad hoc networking capability form a message-passing network on roads in an urban area. The network is infrastructure-less and wireless-connected. The urban road connections in a city are fixed and thus regulate network topology and variations and extensions. Since a vehicle's wireless sensing and transmission range is restricted by its transmitter device and power, and vehicles are free to move independently, the inter-vehicle link is dynamic in frequent setting up and disconnecting, which makes the network a delay tolerant network (DTN) based on message buffering and forwarding for communication. The fundamental routing decision in such network is, for a node buffering messages to decide for each message when encountering a neighbor node, whether to duplicate the message to the other node or not.

2.3 Node Clusters and Message Coverage

As showed in Fig. 1, an urban road in the network can be interpreted as a long tube, so the nodes moving on the same road are naturally grouped as a cluster. Node follows its route and its trajectory on each road are predictable in a short period of time. We propose a very straightforward node clustering method based on roads. Instead of using complex calculation for clustering formation and maintenance, or electing cluster head to coordinate, each node is automatically aware of which cluster it belongs to by only determining which current road it is running on. Nodes are truly independent and equal in a cluster. Unnecessariness of cluster heads also avoid the hidden risk of routing bottleneck and paralysis due to head node malfunction.

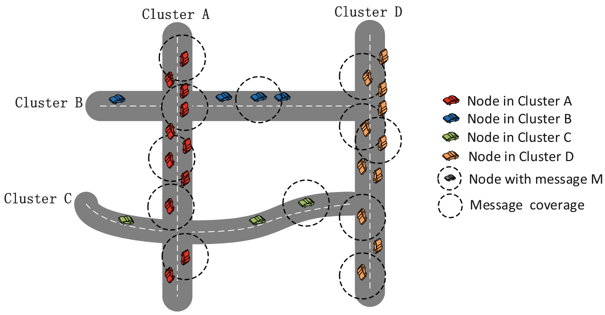


Fig. 1. Road-based node clustering in VANET

For each message being transmitted in the network, there are some vehicular nodes carrying (buffering) it on each node. If we denote such node's wireless communication range as a circle in Fig. 1, then every road is partially covered by such circles, which represent the spread of a specific message in the network, or what we called, the message coverage on each road and in the network. Obviously as soon as the destination node of a given message is under the message's coverage, the message can be successfully delivered. The best coverage is achieved when every node buffers a duplicate of the message, which brings in the fastest message delivery like in Epidemic routing [4], but also has to tolerate the extreme message transmission overhead and consumption of limited buffer space. The proposed routing algorithm in next section is an attempt to trade off for equilibrium of these conflictions.

3 Cluster-Based Routing with Multi-attribute Consideration

In this section, we present a Cluster-based Routing algorithm with Multiple Attributes considered (CR-MA). A message carrier node considers both local message coverage on road and the node's current state to decide whether to delegate the message to an encountered neighbor node. CR-MA optimizes message coverage and overhead when making forwarding decision by predicting the message coverage in each cluster and leveraging four attributes of node, including node velocity, buffer statue, adhesiveness to cluster, and the cluster message coverage.

3.1 Message Coverage Prediction

As defined in Sect. 2.3, we use message coverage to measure the spread of a specific message on a road, i.e. in a cluster. Given M indicating a message, R as a road, the coverage of M on R is computed as $coverage_R^M$ in formula (1). Each message duplicate covers a segment of the road length defined as the diameter of the node's wireless communication range who carries the message. N_M is the predicted number of nodes on road R with message buffering message M.

$$coverage_R^M = \min\left(1, \frac{N_M \cdot 2r}{D_R}\right) \quad (1)$$

r is the wireless communication radius of a node. D_R denotes the length of road R.

On road, every node tracks and predicts message spread by exchanging message buffering information with encountered nodes. The specific procedures that nodes compute a message M's coverage on a road are as follows:

Procedure 1. When node i enters a new road R, it initializes two information tables, one for tracking the prediction of message coverage, abbreviated as *MC*, and the other for remembering the message information of encountered nodes, abbreviated as *NM*. Each entry structure in *MC* is formatted as $\langle M, coverage_R^M \rangle$, where $M \in M_i(M_i$ is the collection of messages buffered by node i. The initial value for $coverage_R^M$ is $2r/D_R$ when only the current node is known to carry message M. An entry in *NM* is formatted as $\langle NID_j, MID_j, v_j, t_j \rangle$, where NID_j indicates the node ID encountered, MID_j the collection of the message IDs carried by node j, v_j the speed of node j no more than the speed limit v_{limit} on road R, and t_j the timestamp of the latest encounter. For the predicted N_M on node i as in formula (1), it is computed as

$$N_M^i = \sum_{j \in NM_M^i} aging^{t-t_j} \cdot \frac{v_j}{v_{limit}} \quad (2)$$

NM_M^i is a subset of *NM* table on node i with all encountered node's IDs which carries message M. *aging* values between 0 and 1 indicating the decay ratio with time elapse. Although the current node is aware that node j carries when last encountered, this awareness becomes less affirmative when time elapses, and if the node's moving velocity is high. t is the current time.

Procedure 2. In two cases a node needs to update its *NM*.

Case 1: When node i encounters node j, they exchange information about each other's id, velocity and the IDs of collection of messages buffered. Then they synchronize *NM*. If a node finds a record of new node in the encountered node's *NM*, it replicates the new record into its own *NM*. They also update the *NM* record of the other node's encounter time.

Case 2: Node i updates its record about node j in *NM* when i relays a message to j.

Procedure 3. In two cases a node adds a new record to *MC*.

Case 1: When node i initiates a new message, it adds a new record with the initial value for message coverage to *MC*.

Case 2: When node i receives a new message, it then calculates the new message's coverage and adds the record to *MC*.

3.2 Multi-attribute Decision-Making

We use multiple attributes to aid on each node's decision of selecting the appropriate relay nodes for messages. Table 1 lists the four aspects of node attributes considered. The forwarding decision for message M from node i to j is computed as a probability, which is a combination of subset probabilities in the four considered aspects of node and network attributes. The following discussion assumes node i come across with j.

Table 1. Subset probabilities for message forwarding

$P_{velocity}$	The forwarding probability based on node speed
P_{buffer}	The forwarding probability based on buffer availability
$P_{vitality}$	The forwarding probability based on node vitality
$P_{coverage}$	The forwarding probability based on message coverage

The forwarding probability based on node speed. Anode tends to spread messages faster to others when moving at higher speed. We compute the probability of message forwarding based on the relation moving speed of the two nodes as follows:

$$P_{velocity} = \begin{cases} \frac{|v_i - v_j|}{v_{limit}}, i \text{ and } j \; move \text{ in the same irection} \\ \frac{v_i + v_j}{2 \cdot v_{limit}}, (i \; and \; j \; move \; in \; the \; opposite \; direction) \end{cases} \quad (3)$$

Where v_i indicates the speed of node i, and v_j indicates the speed of node j.

The forwarding probability based on node buffer availability. To balance the message buffer usage among nodes, the message forwarding decision-making prefers to store messages on those with more buffer space. The probability is computed as

$$P_{buffer} = e^{-\frac{\beta_i}{\beta_j}} \quad (4)$$

Where β_i and β_j are the ratio of free buffer space on node i and j, respectively. $P_{Buffer} = 0$ if the message buffer on node j is full ($\beta_j = 0$).

The forwarding probability function based on the node's vitality. The vitality of a node on a road is the remaining lifetime of this node on the road. Without human input and GPS-based route planning, it is impossible to predict when a vehicular node is leaving the current road. Instead we use the duration of existence on this road and the node's speed as clues to estimate a node's remaining vitality on the current road. The assumption is that the longer a node has been running on a road, and the faster the running speed, the node is more likely to exit. The forwarding probability on node vitality is calculated as follows:

$$P_{vitality} = 1 - e^{-velocity} \cdot e^{-duration} \tag{5}$$

If node i has been moving at speed *velocity* for a *duration* time period.

The forwarding probability based on message coverage on the road. We assume that node i is moving on road R with the message M. We interpret a road as a long tube and the direction and range of a node is limited to the two ends of the road. So M can cover the whole road as long as enough nodes carry copies of M. More duplicates of the message provides better coverage on the road, which brings in higher change of covering the message destination node. However, more redundant message copies also consumes node buffer space, and leads to worse routing cost-effectiveness.

We use the message coverage calculated in formula (1) on current road as a guidance when deciding to delegate a duplicate of each message to the encountered node.

$$P_{coverage} = e^{-coverage_R^M} \tag{6}$$

Multi-attribute Decision-making model. To combine all the attributes together for a probability $P_{i,j}^M$ of node i forwarding the message M to node j, we combine the above four probabilities as

$$P_{i,j}^M = w_1 P_{velocity} + w_2 P_{buffer} + w_3 P_{vitality} + w_4 P_{coverage} \tag{7}$$

Where $\sum_{i=1}^{4} w_i = 1$, $(0 < w_i < 1)$, w_1, w_2, w_3 and w_4 are weights of the node velocity, node buffer availability, node vitality and message coverage in the multi-attribute decision-making model. We choose different weights for diverse network scenarios by using the Analytic Hierarchy Process (AHP) [12]. AHP is a pragmatic structured technique for organizing and analyzing complex decisions, based on a small amount quantitative data. We determine different weights by different judge metrics in various network scenarios using AHP. For example, when the size of node's buffer is low and the cache capacity of node is the key factor restricting the network performance, we can get higher weight of the node's the remaining buffer space by AHP.

3.3 CR-MA Routing Process

As described in Sect. 2.3, every node consciously belongs to a cluster based on its current road location. When two nodes run across each other, there are only two cases of their belonging clusters, either the same one (thus on the same road) or different (in the area of road intersection). Nodes within each other's sensing and communication ranges exchange information about messages buffered, and then decide on delegating messages to the other based on the multiple node and message states and attributes.

(1) **Intra-Cluster Forwarding**

If there are messages whose destination is the neighbor node, the owner node forwards messages preferentially. Otherwise the owner node calculates the forwarding probability $p_{i,j}^M$ of each message regarding to the neighbor node based on the multi-attribute decision-making model. The calculated probability is benchmarked against a configurable threshold γ, and a higher-than-threshold $p_{i,j}^M$ triggers the message replication from node i to j.

(2) **Inter-Cluster Forwarding**

Intuitively nodes running on different roads and meeting each other in intersection areas are occasional, and the connection duration is momentary when both moving in crossing directions. We can use the same multi-attributed based probability $p_{i,j}^M$ for decision of inter-cluster message transmission, except for the message coverage prediction on the destination road is simply set as 0, since the owner node has no clue about the message distribution on a different road. Correspondingly the threshold γ for inter-clustering forwarding is set smaller than in the intra-clustering case.

The choice of threshold value γ should be configured based on the resource available of vehicular nodes, the scale of road map, and the affordable network routing overhead. While a higher threshold reduces the message routing cost, it may also grow the message delivery ratio and timeliness.

4 Experiments and Evaluation

We apply the ONE simulator as the experimental platform and implement the proposed CR-MA routing decision making algorithm as a new routing module in it.

4.1 Simulation Configuration

Table 2 shows a summary of ONE simulation configuration. Based on the original Helsinki road map, we simplify the road map by eliminating and combining trivial roads. The derived road network is composed of 60 major roads. The total length of roads is 128.4 km, individual length between 4325 m and 250 m, as in Fig. 2.

Table 2. Simulation parameters.

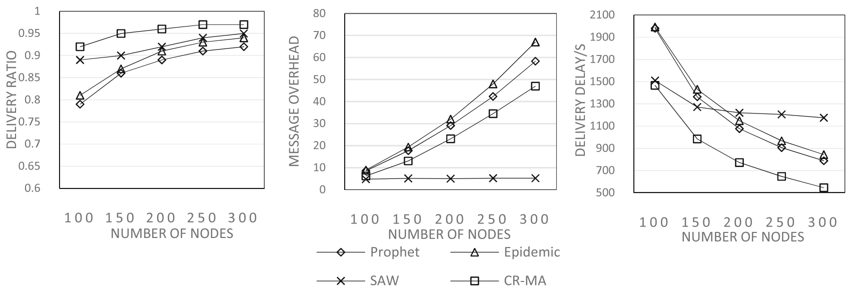
Parameter	Value
Simulation time/h	6
Movement model	Shortest Path Map Based movement
Number of nodes	100, 150, 200, 250, 300
Speed/(m/s)	2.7–13
Buffer size/MB	10, 15, 20, 25, 30
Transmission rate/(kbit/s)	500
Transmission range/m	20
Message sending interval/s	3
Message's TTL/min	300
Message size/kB	500

Fig. 2. The original (left) and modified (right) road network in Helsinki, Finland

To benchmark the proposed CR-MA for performance, we compare it with other well-known high-quality DTN routing algorithms including Epidemic, Prophet and SAW. The value of Prophet's parameters are set as $p_{init} = 0.75$, $\beta = 0.25$, $\gamma = 0.98$. The maximum number of message copies to spray for SAW is 6. The elimination mechanism of redundant messages [13] is applied to each routing algorithm in order to eliminate the message copies after successful delivery.

4.2 Simulation Results and Discussion

The performance of routing algorithm under various node populations. In order to test the scalability of the proposed algorithm, we conduct simulations with node population varying from 100, 150, 200, 250 to 300. Node buffer is fixed as 40 messages.

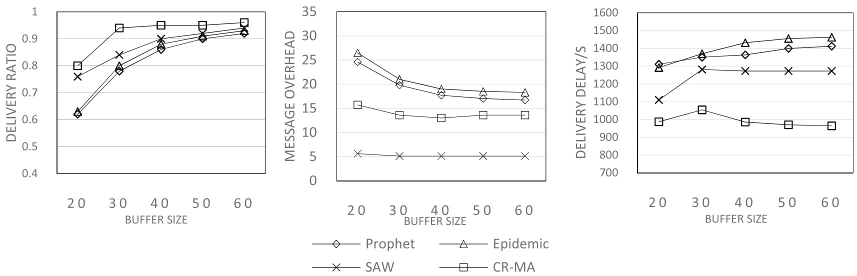


(a) Delivery ratio (b) Message overhead (c) Delivery delay

Fig. 3. Routing performance with various numbers of nodes

Figure 3(a) shows the delivery ratios of all algorithms rises with the growth of the number of nodes. CR-MA and SAW achieve consistently high delivery ratios in small or large populations, with the former always performing the best. This is because CR-MA uses a multi-attribute decision-making model to achieve nearly optimal message forwarding across nodes. Figure 3(b) shows while the message overhead of SAW is the lowest because of its limit on the number of message copies, the message overhead of CR-MA is lower than Epidemic and Prophet. CR-MA avoids unnecessary message flooding by considering the message coverage on each road (cluster). Figure 3(c) shows the delivery delay of all algorithms decreases when the number of nodes grows. This is because messages are forwarded to more nodes and the multipath of a message can reduce message arrival delay. Compared with the others, the delivery delay of CR-MA is the minimum. Messages reach their destinations faster as a result of CR-MA tending to forward messages to faster-moving nodes.

The performance of routing algorithm under various buffer size. Buffer availability is another important attribute in CR-MA affecting forwarding decision making. We also experiment various message buffer sizes with the node number fixed as 150.



(a) Delivery ratio (b) Message overhead (c) Delivery delay

Fig. 4. Routing performance with various buffer sizes

Figure 4(a) shows the delivery ratio of CR-MA is superior to Epidemic, Prophet and SAW. The delivery ratios of all algorithms rise up as the buffer size of nodes increases because a larger buffer space indicates more and longer carrying of messages, which increases the chance of messages being transmitted to their destinations. Figure 4(b) the message overhead of CR-MA is consistently lower than Epidemic and Prophet. Buffer availability is another key attribute considered when forwarding messages. Figure 4(c) shows the delivery delay of all algorithms is increased when the buffer size of node increases because that messages are cached longer. Compared with competitors, CR-MA achieves the shortest delivery delay.

5 Conclusion

This paper proposes a clustering based VDTN routing algorithm with multi-attribute decision-making (CR-MA). CR-MA exploits the layout and organization of roads as restriction of node movement and message spread. Nodes on the same road are organized in a cluster to collaborate on message spreading on this road. We define the concept of message coverage on road to predict and increate the probability of a message reaching its destination. The coverage value, amid other node and network attributes including node velocity, vitality, and availability, help make heuristic decision on when and where to replicate messages. The proposed routing algorithm is verified in simulations and benchmarked with other major algorithm. CR-MA demonstrates improvement and superiority in message delivery ratio, delay and cost. As our future work, we will explore on an automatic configuration for parameter choices so that the proposed CR-MA can be self-adaptive in different application scenarios. We will also try to adopt the current road-based clustering and message coverage mechanism for realistic applications such as collaborative road traffic awareness and forecast.

Acknowledgments. This material is based upon work supported by the China NSF grant No. 61472189, the CERNET Next-Generation Innovation Project under contract No. NGII20160601, and the CASC fund No. F2016020013. Options and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

References

1. Abdelkader, T., Naik, K., Nayak, A., Goel, N.: A performance comparison of delay-tolerant network routing protocols. IEEE Netw. **30**(2), 46–53 (2016)
2. Spyropoulos, T., Psounis, K., Raghavendra, C.S.: Single-copy routing in intermittently connected mobile networks. In: Proceedings of IEEE SECON, pp. 235–244 (2004)
3. Spaho, E., Barolli, L., Kolici, V., Lala, A.: Performance evaluation of different routing protocols in a vehicular delay tolerant network. In: International Conference on Broadband and Wireless Computing, Communication and Applications, pp. 157–162 (2015)
4. Jin, Z., Wang, J., Zhang, S., Shu, Y.: Epidemic-based controlled flooding and adaptive multicast for delay tolerant networks. In: Proceedings of the 2010 Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, pp. 191–194 (2010)
5. Kumar, S., Ahmed, S.H., Qasim, U., Khan, Z.A., Amjad, N., Azeem, M.Q., et al.: Analyzing link and path availability of routing protocols in vehcular ad-hoc networks. J. Basic Appl. Sci. Res. **4**(2), 189–206 (2014)
6. Benamar, N., Singh, K.D., Benamar, M., Ouadghiri, D.E., Bonnin, J.M.: Routing protocols in vehicular delay tolerant networks: a comprehensive survey. Comput. Commun. **48**(8), 141–158 (2014)
7. Ouadrhiri, A., Rahmouni, I., Kamili, M., Berrada, I.: Controlling messages for probabilistic routing protocols in delay-tolerant networks. In: Proceedings of International Symposium on Computers and Communications, pp. 1–6 (2014)
8. Burgess, J., Gallagher, B., Jensen, D., Levine, B.N.: Maxprop: routing for vehicle-based disruption-tolerant networks. In: Proceedings of IEEE INFOCOM, vol. 6, pp. 1–11 (2006)
9. Li, Z., Li, Q., Zhang, H., Liu, F.: Closely social circuit based routing in social delay tolerant networks. J. Comput. Res. Dev. **49**(6), 1185–1195 (2012)
10. Han, J., Shi, J., Ren, Y.: DTN routing algorithm based on region segmentation. Comput. Sci. **42**(10), 113–116 (2015)
11. Wang, E., Yang, Y.J., Li, L.: A clustering routing method based on semi-markov process and path-finding strategy in DTN. Chin. J. Comput. **38**(3), 483–499 (2015)
12. Fan, G., Zhong, D., Yan, F., Yue, P.: A hybrid fuzzy evaluation method for curtain grouting efficiency assessment based on an AHP method extended by D numbers. Expert Syst. Appl. **44**, 289–303 (2016)
13. Small, T., Haas, Z.J.: Resource and performance tradeoffs in delay-tolerant wireless networks. In: ACM Workshop on Delay Tolerant Networking, pp. 260–267 (2005)