OPEN

Contact ability based topology control for predictable delay-tolerant networks

Hongsheng Chen & Chunhui Wu✉

In predictable delay tolerant networks (PDTNs), the network topology is known a priori or can be predicted over time, such as space planet networks and vehicular networks based on public buses or trains. Due to the intermittent connectivity, network partitioning, and long delays in PDTNs, most of the researchers mainly focuses on routing and data access research. However, topology control can improve energy effectiveness and increase the communication capacity, thus how to maintain the dynamic topology of PDTNs becomes crucial. In this paper, a contact ability based topology control method for PDTNs is proposed. First, the contact ability is calculated using our contact ability calculation model, and then the PDTNs is modeled as an undirected weighted contact graph which includes spatial and contact ability information. The topology control problem is defined as constructing a minimum spanning tree (MST) that the contact ability of the MST is maximized. We propose two algorithms based on undirected weighted contact graph to solve the defined problem, and compare them with the latest method in terms of energy cost and contact ability. Extensive simulation experiments demonstrate that the proposed algorithms can guarantee data transmission effectively, and reduce the network energy consumption significantly.

In recent years, Predictable Delay/Disruption Tolerant Networks (PDTNs) have attracted more and more researchers' attention. It is a kind of DTNs^{1,2} whose topology is known a priori or can be predicted over time, such as vehicular networks based on public buses or taxi cabs, pocket switched networks based on human mobility and space satellite networks.

Due to the intermittent connectivity, network partitioning, long delays and node mobility, the PDTNs are not always connected, and the data is delivered in a long delay. The topology dynamically changes over time. Therefore, how to transmit data in each pair nodes successfully is very important. Many routing protocols for PDTNs transmit data from the source to the destination by utilizing opportunistically existing and time varying routing paths. However, the routing protocols only ensure transmit data from the source to the destination using some kinds of redundancy and exploit routing opportunity greedily, which may lead to unnecessary resource consumption.

Generally, topology control can avoid extra resource consumption in data transmitting, which can maintain network connectivity while minimizing the energy consumption and reducing radio interference. It is also true in DTNs, especially for PDTNs³. In PDTNs, a cyclic time C_t exists, and in each C_t , the network topology changes are identical. Pocket switched networks based on human mobility^{4,5}, vehicular networks based on public buses or taxi cabs^{6,7}, disaster-relief networks, and space communication networks^{8–10} are all PDTNs. Thus, it is not necessary to use all possible transmitting opportunities, we can use topology control mechanism to select certain transmitting opportunities to delivery data successfully in such network with lower resource consumption, especially energy consumption and connectivity.

However, it is not feasible to apply the existing topology control mechanism directly to PDTNs because the topology of PDTNs changes dynamically over time. As far as we know, Huang et al.^{11–14} and Chen et al.^{15,16} studied topology control in PDTNs. Yan et al.^{17,18} and Zhang et al.¹⁹ are also studied topology control in PDTNs. Taking into account temporal characteristics, they modeled PDTNs as a directed space–time graph containing spatial and temporal information. In our previously research of the literature¹⁵, we also considered the probabilistic connections of networks. However, the previously researchers of topology control in PDTNs are all not considering the contact opportunity which is an important characteristic in PDTNs.

In this paper, we use the contact opportunity of the nodes of PDTNs in a cycle time to calculate the contact ability, and then construct an undirected weighted contact graph model. The contact opportunity is an important

School of Computer Science and Technology, Hubei University of Science and Technology, Xianning 437100, China.
✉email: chenhs1981@163.com

characteristic of PDTNs, and it is very important for transmit data. A contact ability based topology control mechanism is proposed to satisfy the performance requirement while maximizing the contact ability. Our major contributions are summarized as follows:

1. The contact ability model is constructed.
2. The PDTNs for topology control is modeled as an undirected weighted contact graph. The topology control problem is defined as finding the minimum spanning tree (MST) of such a graph. Two algorithms are proposed to find the minimum spanning tree.
3. The constructed MST can maximize the contact ability of transmitting data with the certain energy consumption.

The rest of this paper is organized as follows. In Section “[Related works](#)”, we summarize related works in topology control in ad hoc networks and PDTNs. In Section “[Node contact opportunity analysis](#)”, contact ability based topology control problem and model are defined. Section “[Contact ability based topology control algorithm](#)” describes the details of the heuristic algorithms proposed in this paper. Section “[Performance evaluation and analysis](#)” presents the simulation results of the proposed algorithms. Finally, Section “[Conclusion](#)” concludes the paper and points out future research directions.

Related works

Contact graph based routing in DTNs. Many researches are focus on contact graph based routing in DTNs. Seguí et al.²⁰ proposed an Enhancing Contact Graph routing of delay-tolerant spatial networks, which focuses on the analysis of contact graph routing (CGR) based on spatial networks. CGR uses the predictability of contacts to make routing decisions. In their simulations, scenes of similar missions to Mars and the Moon were used to collect statistics on routing protocol performance in terms of latency and buffer usage. They improved the basic cost function of CGR to avoid routing loops and suggested using Dijkstra’s shortest path algorithm for path selection. Bezirgiannidis et al.²¹ proposed CGR with the earliest transmission opportunity (CGR-eto) and overbooking management. While evaluating the enhancement effect of the original CGR, they also introduced an experimental version of CGR-eto, which uses the information of local routing data to calculate the queuing delay of all hops that reach the destination through the path, not just the first jump. Caini and Firrincieli²² studied the applicability of contact graph routing (CGR) in LEO satellite DTN communication, focusing on two practical application scenarios of earth observation and data mule. The results obtained by running ION, the DTN Bundle protocol and the Linux test bed realized by CGR developed by NASA highlight the advantages of CGR applied to low-orbit satellite communications. Araniti et al.²³ investigate the Contact Graph Routing in DTN Space Networks, especially the applicability and performance of the DTN protocol stack, and evaluate the experimental experience of the CGR presenting results. Fraire and Finochietto²⁴ proposed the challenge of the connection plan of the anti-interruption satellite network. They discussed the increasing complexity of considering routing and traffic information to enrich the planning process, and therefore the need to implement a contact plan calculation element to support spatial DTN operations.

All these contact graph based routing are also focusing on delivering data from source to destination. They do not consider the topology control.

Topology control for PDTNs. There are few studies on topology control under PDTNs, as far as we know, Huang et al.^{11–14}, Chen et al.^{15,16}, Yan et al.^{17,18} and Zhang et al.¹⁹ studied the topology control of PDTNs. Huang et al. modeled the time-evolving network as a weighted space–time graph, which include temporal and spatial information. They proposed an cost efficient topology design (CETD) for PDTNs, and construct a sparse network that meets two goals, that is, every pair of nodes can be connected and the total energy consumption is the smallest. Moreover, the unreliable links is considered in literature¹³. The effectiveness of above proposed topology control methods are all verified by simulation. Yan et al.^{17,18} also modeled the PDTNs as a directed space–time graph that includes spatial and temporal information. They proposed a topology control strategy by using a kth shortest paths algorithm. Zhang et al.¹⁹ investigated the topology control problem in spacecraft networks where the time-varying topology can be predicted, and they modeled the spacecraft networks as a directed space–time graph, which also includes both temporal and spatial topology information too. They proposed five heuristic algorithms, which can significantly maintain low topology cost efficiency ratio while achieving high reliable connectivity. Although above studies are all construct a sparse network to ensure that each pair of nodes can communicate, but they only consider the spatial and temporal information. The contact opportunity that is an important characteristic in PDTNs is not considered.

To our best knowledge, there are no previous results on topology control based on contact opportunity of PDTNs. This paper is the first attempt to study topology control for PDTNs based on contact ability through constructing MST. We believe that topology can be controlled more wisely and efficiently if the network evolution over time is known.

Node contact opportunity analysis

In DTNs, the data transmission is achieved through opportunistic contact caused by node motion. Thus, how to define the node contact ability is crucial. In this paper, a network model based on contact graph is established, and the process of node contact with each other under different conditions is analyzed. The contact process of node in single hop and multi-hop is described, and the calculation method of contact probability of each pair of nodes is described. Moreover, the node contact ability measurement is defined.

Figure 1. DTN network partitions.

Figure 2. Node contact time.

Network model. The contact of the node is described by the network contact diagram $G(V, E)$. In the contact graph, V is the set of the nodes, E is the set of the edges. The edge e_{ij} represent node i and j can contact. Similar to the literature^{25,26}, it is assumed that the contact interval between each pair of nodes follows the exponential distribution of the parameter λ_{ij} , so the contact process between nodes i and j can be described as the Poisson process with contact rate λ_{ij} , λ_{ij} can be calculated by real-time time-average. Although the literature²⁵ shows that the node contact time interval is composed of the power-law distribution and the exponential distribution, there is no agreement on the distribution of the contact time of the pair of nodes. The exponential model used is validated in the actual DTN environment given in^{27–29}.

Figure 1 shows an example of a typical DTN network. In DTN networks, due to the randomness of nodes, the links between nodes in a large range are unreliable and intermittent. Due to the nodes movement, the network will be split into many smaller network partitions.

Figure 1 shows the three network partitions 1, 2 and 3, which is circled with virtual line. Within the network partition, the circle with the letter number represents the mobile node, and the solid line of the connected node represents the contact process between the nodes. The dotted line with the connected node represents the node contact process after a period of time. For example, the nodes L, M, N, O, and P belong to the network partition 2, the nodes L and M are contact with λ_{LM} contact rate. After a while, node M will leave the network partition 2 and enter the network partition 3, and contact with J by λ_{MJ} contact rate.

Network information. In order to effectively transmit data, it is very important to obtain the network information when the nodes contact. The information include (1) the contact interval between nodes, (2) the node set in contact with the node, (3) the duration of contact between nodes. Due to the instability of the connection in DTNs, it is impossible to share these network information in a global scope, so that the optimal data transmission path cannot be given in the whole network. Thus, exploiting node's local network information to choose the appropriate path to effectively transmit data.

(1) Single-hop network information maintenance

In the network, each node is identified by a unique ID. When node i contact with node j , the node i will record the node number (here is j), the contact time, and the contact duration, and then put it as a record into their own contact information table. Thus, after a period of time, the contact information table of the node i records its contact history information, and the contact history shown in Fig. 2 can be expressed as the contact table shown in Table 1.

(2) Multi-hop network information maintenance

From the perspective of single-hop to study the contact of the node cannot fully tap the node's mobility and data transmission capacity. For example, node i only contact nodes A and B in a period of time, but node A can contact nodes C, D, and E in the future. When node j needs to send data to E, in the case of single-hop contact, it will not forward data to node i after contact node i , instead continue to wait for the chance of contact with E. Thus, the opportunity to forward data to i via A to E is lost, and reducing the possibility of successful data

Node number	The contact time	Contact duration
N1	T1	L1
N2	T2	L2
N3	T3	L3
N1	T4	L4
N4	T5	L5
N5	T6	L6
N3	T7	L7
N3	T8	L8

Table 1. Node contact information table.

The source node	Target node	Contact rate	Update time
N1	N2	λ_1	5
N1	N3	λ_2	9
N2	N4	λ_3	13
N2	N3	λ_4	20
N4	N7	λ_5	30
N7	N9	λ_6	54

Table 2. Contact rate estimation table of node i.

The source node	Target node	Contact rate	Update time
N1	N3	λ_7	5
N4	N6	λ_8	7
N2	N3	λ_9	13
N9	N2	λ_{10}	20
j	N2	λ_{11}	30
j	N11	λ_{12}	54

Table 3. Contact rate estimation table of node j.

forwarding, increased access delay. For this reason, it is necessary to consider the contact ability of the node in the multi-hop range.

In the process of maintaining the single-hop network information, the node maintains the multi-hop network information by exchanging the additional network information during the contact process. This information mainly includes the estimated value of the contact rate between the nodes that contact with it and the other nodes. Since the contact condition of the node in DTNs is constantly changing, this information is local, which is called the contact rate estimate (described by Section “Node contact probability(1)”). In this way, each node will maintain two tables, the node contact table and the contact rate estimate table. For a contact rate estimate, all records associated with that node in the table should be updated when the maintenance node contact to the nodes in the table again. That is because the records in the contact rate estimates have expired and need to be updated with the new contact rate estimates to estimate the network contact as accurately as possible.

The following tables are an example of multi-hop network information maintained. Table 2 shows the current node i contact rate estimation table, Table 3 is the contact rate estimation table of node j.

Node i contact with node j at time 70. First, they add a new contact record between i and j on their respective contact information tables and recalculate the contact rate λ_{ij} between i and j using the method described in Section “Node contact probability(1)” and update the λ_{ij} to the contact rate estimate table, the update time is set to 70. Then, node i and j updates the public records in their contact rate estimates and adds new records. The updated results are shown in Tables 4 and 5.

Tables 2 and 3 are only part of i and j's respective contact rate estimates, after i and j contact, they will update the common record (here the second and fourth rows of Table 2, and update the contact rate of N1, N3 in Table 3 to λ_2 according to the update time of these common records, the contact rate of N2 and N4 in Table 2 is updated to λ_9 and the update time is modified. Next, insert the remaining records in Table 3 into Table 2, and insert the remaining records in Table 2 into Table 3. Finally, the contact rates between i and j are inserted in Tables 2 and 3 and the update times are set to their contact time 70.

The source node	Target node	Contact rate	Update time
N1	N2	λ_1	5
N1	N3	λ_2	9
N2	N3	λ_3	13
N2	N4	λ_9	25
N4	N7	λ_5	30
N7	N9	λ_6	54
N4	N6	λ_8	7
N9	N2	λ_{10}	30
j	N2	λ_{11}	46
j	N11	λ_{12}	60
i	j	λ_{ij}	70

Table 4. The updated contact rate estimation table of node i.

The source node	Target node	Contact rate	Update time
N1	N3	λ_2	9
N4	N6	λ_8	7
N2	N4	λ_9	25
N9	N2	λ_{10}	30
j	N2	λ_{11}	46
J	N11	λ_{12}	60
N1	N2	λ_1	5
N2	N4	λ_4	13
N4	N7	λ_5	30
N7	N9	λ_6	54
i	j	λ_{ij}	70

Table 5. The updated contact rate estimation table of node j.

Figure 3. The network contact graph G_i of node i.

After exchanging network information between node i and j, the local contact map G_i is established according to their own maintenance of the network information. Figure 3 shows a schematic representation of G_i .

Node contact probability. In this paper, the contact ability of nodes is measured through the probability ways. As can be seen from the network model in Section “Network model”, it is assumed that the contact of each pair of nodes is subject to the Poisson process. The node i calculates the contact rate λ_{ij} with the node j based on the network information maintained by itself. In this section, the contact rate between nodes is considered from the perspective of single-hop and multi-hop.

Figure 4. Opportunity path diagram.

(1) Calculating inter-node contact rate

As described in Section “Network information(1)”, each node maintains a contact information table that records the contact history information of the node, including the contact node ID, contact duration, and so on. In this paper, it is assumed that the contact between nodes obeys the Poisson process and the respective contact rates need to be calculated for each node in the contact information table. For node i , the specific method is as follows:

1. The node numbers in the contact information table are merged into the set S .
2. For each of the set numbers s in the collection S , filter out all the records of the node s in the contact table, and define these records as recordset R_s , which size is L_s .
3. Find the maximum and minimum values of the contact time columns in R_s , denote $T_{s\max}$ and $T_{s\min}$ respectively, find the difference $Trange$ of $T_{s\max}$ and $T_{s\min}$, the result of dividing $Trange$ by $L_s - 1$ takes the reciprocal to obtain the contact rate λ is of node i and s .

$$\lambda_{is} = \frac{L_s - 1}{T_{s\max} - T_{s\min}} \quad (1)$$

4. Repeat the process step 2) until all the nodes in S executed in step 2) operation.

(2) Single-hop node contact probability

After knowing the contact rate λ_{ij} between the nodes i and j , the contact probability between single-hop nodes can be calculated. According to the foregoing assumptions, the contact interval between i and j obeys the exponential distribution of the parameter λ_{ij} . Assuming that the current time is t , the lifetime of the data copy d is T (from 0 onwards). From the probability density function of the exponential distribution, we can see that the probability of i and j before T is the following formula.

$$P(X \leq T_d) = 1 - e^{-\lambda_{ij}(T_d - t)} \quad (2)$$

where X is the next contact time of i and j .

(3) Multi-hop node contact probability

Definition 1 The r -hop opportunistic path $PAB = (VP, EP)$ between A and B , which contains the node set $VP = (A, N1, N2, \dots, Nr-1, B)$ and the edge set $EP = (e1, e2, \dots, er)$, the weights are $(\lambda1, \dots, \lambda_r)$.

At the same time, the path weight is defined as the probability that the data is forwarded from A to B along the PAB within time T .

As shown in Fig. 4, the contact interval X_k of the nodes N_k and $N_k + 1$ obeys the exponential distribution of the parameter λ_k . Whereby the total time required for forwarding the data from A to B is $Y = \sum_{k=1}^r X_k$ obey the super exponential distribution³⁰.

Lemma 1³¹ The probability density function (PDF) for Y is a r -hop path of $\lambda1, \lambda2, \dots, \lambda_r$ for the edge.

$$p_Y(x) = \sum_{k=1}^r C_k^{(r)} p_{X_k}(x) \quad (3)$$

$$\text{among them, coefficient } C_k^{(r)} = \prod_{s=1, s \neq k}^r \frac{\lambda_s}{\lambda_s - \lambda_k}.$$

From Definition 1, the path weights can be written as the following formula.

$$p_{AB}(T) = \int_0^T p_Y(x) dx = \sum_{k=1}^r C_k^{(r)} \cdot (1 - e^{-\lambda_k T}) \quad (4)$$

Figure 5. The probability of contact in different situations.

(4) Measurement of contact probability between nodes

With the above description of single-hop contact probability and multi-hop contact probability, it is possible to further analyze the possible contact patterns between nodes.

Figure 5 shows the typical ways in which three nodes are exposed. Among them,

1. Figure 5a describes a simple single-hop contact, that is, the node i, j without the help of other nodes in the case of independent contact with each other.
2. Figure 5b describes the case where the nodes i, j are contacted by means of an opportunistic path. In these cases, the contact between i and j requires a series of intermediate nodes, and the contact of these intermediate nodes directly affects i, j 's contact probability.
3. Figure 5c describes a case where (a) and (b) are mixed, where there is a direct contact between i and j , with the presence of one or more jumper paths, which makes i, j probability is greater than the probability of direct contact or through the opportunistic path, but the probability calculation in this case becomes complicated. In the initial stage of the network, the information of the network is less, the difficulty of the calculation of the contact probability will be relatively small, but with the continuous accumulation of network information, the complexity of the network will continue to increase, the corresponding calculation will increase the difficulty;
4. The situation in Fig. 5d is more complicated. In this case, there are multi-hop opportunistic paths between nodes i and j and there are common nodes on these paths, which have a very large effect on the whole contact probability. When a public node disappears for some reason in the network, the edges of the network contact graph associated with the node will disappear and the corresponding contact map between i and j will change dramatically. The update of the network information cannot be quickly broadcast in the network, then Fig. 5d in this case the probability of the calculation will be very large deviation.

Through above analysis, we can know that the probability of contact between nodes is different in different situations. In this paper, for the sake of simplicity, this paper does not attempt to calculate the node contact probability in the case of Fig. 5c,d. The Dijkstra algorithm is used to derive the opportunity path with the shortest number of hops, and the basic probability of the path is used as the contact probability between nodes.

Contact ability calculation. For a given node, its ability to reach other nodes is one of the most important features in the DTN network.

Researchers have proposed a number of ways to tap the node's ability to contact, these methods are centralized, but also distributed. The centralized method assumes that the nodes in the network know the global information of the network. Using this information, the nodes judge the strength of their contact ability in the whole network and pass the data to the nodes with the best or relatively good contact ability. The advantage of this

Figure 6. Undirected weighted contact graph.

approach is the decision-making accuracy, good performance, the disadvantage is the need for additional means in the whole network broadcast network information. In the distributed method, the node uses the historical contact opportunity to exchange the network information, through the existing network information processing to estimate the network, calculate their own contact ability and in contact with other nodes in contact with other contact ability, thereby deciding whether to forward data.

As described in Sects. Section “Network information(1)” and “Network information(2)”, each node maintains a node contact table and a contact rate estimate. According to the method of Sections “Node contact probability(1)” and “Node contact probability(2)”, the contact information table can be used to calculate the probability of single-hop contact between nodes. According to the description of Section “Node contact probability(4)”, the contact rate estimation table can be used to select the opportunity path between nodes to calculate the multi-hop contact probability. In this paper, we discuss the probability of single-hop and multi-hop contact as the probability of contact between nodes.

Definition 2 The node i contact ability measure C_i is described by the following formula.

$$C_i(T_d - t) = \sum_{j \in N_i} p_{ij}(T_d - t) \quad (5)$$

where T_d is the lifetime of the data copy d from time 0, t is the current time, and N_i is the set of nodes that can be reached. $p_{ij}(T_d - t)$ is the probability of the contact of i and j in the time $T_d - t$ from the time t , and the calculation method has been described above.

From the above equation, $C_i(T_d - t)$ calculates the expected value of the number of nodes that the node i may be exposed to in future $T_d - t$. The expected value reflects the ability of node i to contact the node during the remaining lifetime of replica d .

Undirected weighted contact graph model. Firstly, according above formula 5, the contact ability of node i can be counted out at a cycle time, where T_d is the lifetime of the data replication d from time 0, t is the current time, and N_i is the set of nodes that can be reached. In this paper, we set the current time is 0 and T_d is a cycle time of PDTNs. $p_{ij}(T_d - t)$ is the probability of the contact of i and j in the time $T_d - t$ from the time t .

And then average the contact ability of each pair of nodes contact with each other in the networks.

Finally, according the above calculating contact ability of each pair of nodes, one undirected weighted contact graph is constructed. Such as the following Fig. 6, the vertex represent the nodes of the PDTNs, the edge of the nodes represent the two nodes can be contact, and the weight of the edge is the reciprocal of the average contact ability of the two nodes.

Thus, the total contact ability of the MST is represented by $C(C)$ shown in the following formula:

$$C(C) = \sum_{K=1}^{N-1} \frac{1}{W(V_i V_j)}$$

($V_i V_j \in G$, $V_i V_j$ is in the minimum spanning tree, $W(V_i V_j)$ is the average contact ability of node i and j , and N is the total number of the vertices of G).

Energy model. In PDTNs, data transmission requires energy, and energy cost is a very important issue. Minimizing the energy cost is the important goal of topology control, thus an energy consumption model needs to be established. Generally, the free space propagation model $P_t(d) \propto d^2$ is widely used. The P_t is the power of the transmitted signal and d is the Euclidean distance between the transmitter and the receiver. In order to simplify the implementation, according to the distance between each pair of nodes, the energy consumption value of an appropriate edge is given from 1 to 5. We use $C(V_i V_j)$ to represent the energy cost of communication between node i and node j , and the total energy cost of MST is represented by $C(G)$, as shown in the following formula:

$$C(G) = \sum_{k=1}^{N-1} C(V_i V_j)$$

($V_i V_j \in G$, $V_i V_j$ is in the minimum spanning tree and N is the total number of the vertices of G).

Problem statement. The aim of topology is to maintain network connectivity with the minimum energy cost. In this paper we mainly focus on the nodes can transmission each other efficiently in the maximum contact ability in PDTNs. Topology control problem can be defined as a tree construction problem, which is finding the minimum spanning tree of undirected weighted contact graph model, such that 1) include all nodes of the network; 2) the contact ability of MST is maximized; and 3) the total energy cost of the MST is minimized. Given a Undirected graph $G(V, E, e)$, the problem is to find a MST of $G(V, E, e)$ that satisfies,

$$\begin{aligned} &\text{maximize: } C(C) \\ &\text{subject to: } C(G) \text{ is minimize} \end{aligned}$$

Contact ability based topology control algorithm

Variant Kruskal Algorithm base on Undirected Weighted Contact Graph (VKAUWCG). The basic idea of the first algorithm is finding the biggest contact ability edges that satisfy the energy cost is minimize to ensure all nodes of the networks can transmission each other efficiently.

The VKAUWCG contains the following steps:

Step 1: All the edges are sorted in the ascending order according to the reciprocal of the average contact ability of each pair of nodes defined in the undirected weighted contact graph. And every vertex of the entire network is initiated as a connected component.

Step 2: Choosing the undirected edge with the smallest weighted edge, if the smallest weighted edge is more than one, then select the smallest energy cost edge, and merge the vertexes of this edge as one connected component.

Step 3: Choosing the undirected edge with the second smallest weighted edge. If the second smallest weighted edge is more than one, then select the smallest energy cost edge. If the vertexes of this edge satisfy the certain conditions, they will be merged into the previous connected component. These conditions are:

- (1) The vertexes are in different connected component.
- (2) The vertexes haven't connected with the edges.

Step 4: If the above conditions are not satisfied, the undirected edge with the i th ($i = 3, \dots$) smallest weighted edge is selected and Step 3 is performed repeatedly until all vertexes have been reached by the selected edges. If one connected component contains all these vertexes, this connected component is the Minimum Spanning Tree (MST). Otherwise, the network is not complete connected.

Step 5: The contact ability and the ratio of total cost are calculated. And this algorithm ends with the established MST.

The pseudo code of VKAUWCG algorithm is as follows, in which CC represents the connected component, Ne represents the number of the edge, $V(E_i)$ represent the two vertexes of edge E_i , n represents the number of the vertex:

Input:

The original undirected weighted contact graph, including the energy cost of each edge, the contact ability of each pair of nodes.

Output:

The contact ability, the ratio of total cost and the Minimum Spanning Tree(MST);

```

1: Sort the edges of G in ascending order according to the
   reciprocal of the average contact ability of each pair of
   nodes;
2: k=1;
3: j=k, initiated every vertex as a CC;
4: for i= j;j<Ne;j++ do
5: Select the edges with the ith smallest weight,
   If (If the smallest weighted edge is more than one)
       then
       select the smallest energy cost edge;
       merge the two vertex of this edge as one CC;
6: if (the two vertexes are in different CC) and (the
   vertexes haven't connected with any edges)
       then
7: merge(CC,V(Ei));
8: if all vertexes are connected with n-1 edges and in one
   CC then
9:   MST is constructed;
10: else
11: if all vertexes are connected with n-1 edges, but not
   in one CC then
12: The network is not complete Connected;
13: end if
14: end if
15: end if
16: end if
17: end for
18: Calculating contact ability and the ratio of total cost;

```

The time complexity of line 1 is $O(n^2)$, the time complexity of line 4–16 is $O(n \times \log N_e)$. So the time complexity of the algorithm is $O(n^2 + n \times \log N_e)$.

Theorem 1 *The time complexity of the above algorithm is $O(n^2 + n \times \log N_e)$, in which n represents the number of the nodes in the network, N_e represents the number of the edges of the network.*

Variant Prim algorithm base on Undirected Weighted Contact Graph (VPAUWCG). The basic idea of the second algorithm is to find the biggest contact ability edges that satisfy the energy cost is minimize to ensure all nodes of the networks can transmission each other efficiently by adding nodes. The VPAUWCG contains the following steps:

Step 1: The connected component (CC) is initiated as empty. One vertex is randomly chosen and added to the CC.

Step 2: Choosing the undirected edge with the smallest reciprocal of the average contact ability between nodes connected to the CC's vertex. And to ensure that no closed cycle will form after joining this edge.

Step 3: If the smallest weighted edge is more than one, then select the smallest energy cost edge. The connected vertex of the selected edge is added to the CC if it is not in the CC.

Step 4: If CC still does not contain all the vertexes, goto Step 2 until all vertexes are in the CC. A MST is constructed. Otherwise, the network is not complete connected. Goto Step 1.

Step 5: The contact ability and the ratio of total cost are calculated. And this algorithm ends with the established MST.

The pseudo code of VPAUWCG algorithm is as follows:

Input:

The original undirected weighted contact graph, including the energy cost of each edge, the contact ability of each pair of nodes

Output:

The contact ability, the ratio of total cost and the Minimum Spanning Tree(MST);

```

1: j=1;
2: CC={Vs}; // Vs is a random vertex of the graph
3: i=j;
4: Choosing the undirected edge with the ith smallest
reciprocal of the average contact ability between nodes
connected to the CC's vertex;
5: If (the smallest weighted edge is more than one)
  then
    select the smallest energy cost edge;
6: end if
7: if (the starting vertexes of the edge is not in CC and no
closed cycle is formed)
  then
    8: add them to CC;
9: end if
10: if (CC still does not contain all the vertexes)
  then
    11: Goto 3
  Else
    12: Goto 2
13: end if
14: Calculating contact ability and the ratio of total cost;

```

The time complexity of line 4 is $O(N_e \times \log n)$, and for line 5–13 is $O(n \times N_e \times \log n)$. So the time complexity of the algorithm is $O(n \times N_e \times \log n)$.

Theorem 2 *The time complexity of the above algorithm is $O(n \times N_e \times \log n)$, in which n represents the number of the node in the network, and N_e represents the number of the edges of the network.*

Performance evaluation and analysis

In this section, we will evaluate and analyze our proposed algorithms, namely, Variant Kruskal Algorithm base on Undirected Weighted Contact Graph (VKAUWCG) and Variant Prim algorithm base on Undirected Weighted Contact Graph (VPAUWCG) by comparing their performance in terms of the energy cost and contact ability. We compared our algorithms with GrdLCP¹⁴ and our previous MSTRAG¹⁵ and VPASPG¹⁶. We implement all these algorithms in a simulated environment. The underlying PDTNs are randomly generated from a random graph model or directly extracted from Cambridge Hagggle tracing data³².

In simulation, we mainly consider the following performance metrics and parameters:

1. The ratio of total cost: the performance metric is mainly used to evaluate whether the algorithm is good or bad by the energy efficiency, the smaller the total energy cost, the algorithm is more optimal.
The ratio of total cost = total energy cost value/the whole network energy value.
2. Link density (ρ): It refers to the sparse/dense degree of the network topology.
Link density = the number of links of current topology/ the number of links of the corresponding complete graph of the current topology.
3. Contact ability: It refers to the contact ability if each pair of nodes. In this paper the contact ability can be obtained from formula 5.

Simulations on random networks. We use the undirected weighted contact graph to represent the predictable DTNs. It has 10 nodes and the distribution of nodes is different for each run. The cost of each edge is randomly chosen from 1 to 5, and the reciprocal of the contact ability of each edge is randomly chosen from 0 to 1. For all the simulations, we repeat the experiment for 100,000 times and report the average values of the metric.

Figure 7 depict the impact of link density ρ to the ratio of total cost. As can be seen from the figures, our two algorithms proposed in this paper are always better than the other algorithms, especially when the link density

Figure 7. Ratio of total cost of different link density ρ .

Figure 8. The reciprocal of the average contact ability of different link density ρ .

is smaller. However, with the link density increasing VPASPG and MSTRAG proposed in our previous research is very similar to our two algorithms proposed in this paper. That is because the methods proposed in our paper are constructing a tree for topology control. From Fig. 7 we can also see the ratio of total cost decreases with ρ increasing. Even when ρ is 0.1, the ratio of total cost is in the range of 0.063~0.07, which indicates that even in the sparse network especially in the challenge environment our algorithms could save more energy.

Figure 8 depict the impact of link density ρ to the reciprocal of the average contact ability. As can be seen from the figures. The tendency of the reciprocal of the average contact ability of the two algorithms proposed in this paper decreases with ρ increasing, which indicated that the contact ability of the network is increasing with ρ increasing. That is because with the link density increasing, the density of the networks is very high, thus the contact ability is very big. From Fig. 8 we can also see the reciprocal of the average contact ability is very low when ρ is small. Even when ρ is 0.1, the reciprocal of the average contact ability is in the range of 9~11, which indicates that even in the sparse network especially in the challenge environment our algorithms could ensure the data can be transmitted to the other nodes efficiently. That's because in our simulation the network has 10 nodes. Figure 9 also depict VKAUWCG is better than VPAUWCG when ρ is smaller than 0.3. Thus, if the required contact ability is large, we can use the VKAUWCG, otherwise if the required energy cost is small, VPAUWCG will be used.

Figure 9. The constructed original graph of the Tracing Data using our model.

Figure 10. The constructed graph using the contact ability based topology control methods.

Simulations on tracing data. In this paper, we use the data sets from Cambridge Hagggle data³¹. In this data set, connections among 78 mobile iMote Bluetooth nodes carried by researchers and additional 20 stationary nodes are recorded over four days during IEEE Infocom 2006. In our simulation, we only consider the 78 mobile nodes, if there is a contact between each pair of two nodes, an edge is added to the graph. The weight of the edge is represented by the reciprocal of the sum of the number of contacts of the two nodes. The following Fig. 9 is the constructed original graph of the Tracing Data using our model.

From the original graph of the network, we can see the graph is very dense. If we don't topology control for it, the energy cost will be very large, and the signal interference will be strong. Thus topology control is very necessary.

Figure 10 is the constructed graph using the contact ability based topology control methods proposed in this paper. From the results, we can see the method proposed in this paper is very efficient in terms of energy cost and contact ability. The graph become very sparse, the contact ability is very high, and the energy cost is very small. Due to the links of the network is very less, thus the signal interference is also very small.

Conclusion

In this paper, we model PDTNs as an undirected weighted contact graph which includes spatial and contact ability information and defined the topology control problem as finding a Minimum Spanning Tree(MST) from the graph model to maximize the contact ability. We propose two heuristic topology control algorithm, VKAUWCG and VPAUWCG to solve this problem.

The next step is to consider the following problems: (1) Solving the topology control problem for unpredictable DTNs; (2) Devising a method that can get contact information accurately; (3) Dealing with topology control of PDTNs when some node may fail in the movement; (4) Balancing energy cost of the network.

Received: 23 June 2021; Accepted: 3 November 2021

Published online: 19 November 2021

References

- Juang, P. *et al.* Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebranet. *ACM SIGOPS Oper. Syst. Rev.* **36**(5), 96–107 (2002).
- Wang, Y., Dang, H. & Wu, H. A survey on analytic studies of delay-tolerant mobile sensor networks: Research articles. *Wirel. Comm. Mobile Comput.* **7**(10), 1197–1208 (2007).
- Shao, Y. & Wu, J. Understanding the tolerance of dynamic networks: A routing-oriented approach. In *Proceedings of International Conference on Distributed Computing Systems Workshops (ICDCS)*, 180–185 (2008).
- Hui, P. & Crowcroft, J. Predictability of human mobility and its impact on forwarding. In: *Proceedings of International Conference on Communications and Networking in China*, 543–547 (2008).
- Hui, P., Crowcroft, J. & Yoneki, E. BUBBLE rap: Social-based forwarding in delay-tolerant networks. *IEEE Trans. Mob. Comput.* **10**(11), 1576–1589 (2011).
- Zhang, X., Kurose, J., Levine, B.N., Towsley, D. & Zhang, H. Study of a bus-based disruption-tolerant network: Mobility modeling and impact on routing. In *Proceedings of ACM MobiCom* (2007).
- Piorkowski, M., Sarafijanovic-Djukic, N. & Grossglauser, M. A parsimonious model of mobile partitioned networks with clustering. In: *Proceedings of First Int'l Conf. Comm. Systems and Networks and Workshops (COMSNETS'09)* (2009).
- Merugu, S., Ammar, M.H. & Zegura, E.W. Routing in space and time in networks with predictable mobility, CC Technical Report; GIT-CC-04-07 (2004). <http://hdl.handle.net/1853/6492>
- De Cola, T., Marchese, M. & Raviola, A. power and bandwidth effective data communications in disaster relief operations through a satellite-based disruption tolerant network paradigm. In *Proceedings of IEEE International Conference on Communications (ICC)*, 1876–1880 (2008).
- Burleigh, S. & Hooke, A. Delay-tolerant networking: An approach to interplanetary internet. *IEEE Comm. Mag.* **41**(6), 128–136 (2003).
- Huang, M., Chen, S., Zhu, Y., Xu, B. & Wang, Y. Topology control for time-evolving and predictable delay-tolerant networks. In *Proceedings of IEEE Mobile Ad-Hoc and Sensor Systems* (2011).
- Huang, M., Chen, S., Zhu, Y., Xu, B. & Wang, Y. Cost-efficient topology design problem in time-evolving delay-tolerant networks. In *Proceeding of IEEE Globecom* (2010).
- Huang, M., Chen, S., Li, F. & Wang, Y. Topology design in time-evolving delay-tolerant networks with unreliable links. In *Proceeding of IEEE Globecom* (2012).
- Huang, M., Chen, S., Zhu, Y. & Wang, Y. Topology control for time-evolving and predictable delay-tolerant networks. *IEEE Trans. Comput.* **62**(11), 2308–2321 (2013).
- Chen, H. & Shi, K. Topology control for predictable delay-tolerant networks based on probability. *Ad Hoc Netw.* **24**, 147–159 (2015).
- Chen, H., Shi, K. & Wu, C. Spanning tree based topology control for data collecting in predictable delay-tolerant networks. *Ad Hoc Netw.* **46**, 48–60 (2016).
- Yan, D. *et al.* A topology control strategy with efficient path for predictable delay-tolerant networks. *IEICE Trans. Commun.* **E102.B**(12), 2183–2198 (2019).
- Yan, D., Liu, C., You, P. *et al.* Efficient topology design for load balancing in time-evolving delay-tolerant networks. In *2018 8th International Conference on Electronics Information and Emergency Communication (ICEIEC)* (2018).
- Zhang, W., Ma, H., Wu, T., Shi, X. & Jiao, Y. Efficient topology control for timevarying spacecraft networks with unreliable links. *Int. J. Distrib. Sensor Netw.* **15**(9), 1–18 (2019).
- Seguí, J., Jennings, E. & Burleigh, S. Enhancing contact graph routing for delay tolerant space networking. In *IEEE Globecom*, 1–6 (2011).
- Bezirgiannidis, N., Caini, C. & Tsaoussidis, V. Analysis of contact graph routing enhancements for DTN space communications. *Int. J. Satell. Commun. Network* **34**, 695–709 (2016).
- Caini, C. & Firrincieli, R. Application of contact graph routing to LEO satellite DTN communications. In *IEEE ICC 2012—Selected Areas in Communications Symposium*, 3301–3305.
- Araniti, G. *et al.* Contact graph routing in DTN space networks: Overview, enhancements and performance. *IEEE Commun. Mag.* **53**(3), 38–46 (2015).
- Fraire, J. A. & Finochietto, J. M. Design challenges in contact plans for disruption-tolerant satellite networks. *IEEE Commun. Mag.* **53**(5), 163–169 (2015).
- Balasubramanian, A., Levine, B. & Venkataramani, A. DTN routing as a resource allocation problem. In *Proceedings of the 2007 ACM International Conference on Computer Communication*, 373–384 (2007).
- Zhu, H., Fu, L., Xue, G. *et al.* Recognizing exponential inter-contact time in VANETs. In *Proceedings of IEEE INFOCOM*, 1–5 (2010).
- Conan, V., Leguay, J. & Friedman T. Characterizing pairwise inter-contact patterns in delay tolerant networks. In *Proceedings of the 1st International Conference on Autonomic Computing and Communication systems*, 19 (2007).
- Ioannidis, S., Chaintreau, A. & Massoulié, L. Optimal and scalable distribution of content updates over a mobile social network. In *Proceedings of IEEE INFOCOM*, 1422–1430 (2009).
- Gao, W., Li, Q., Zhao, B. *et al.* Multicasting in delay tolerant networks: A social network perspective. In *Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 299–308 (2009).
- Karagiannis, T., Le Boudec, J. Y. & Vojnovic, M. Power law and exponential decay of intercontact times between mobile devices. *Mobile Comput.* **9**(10), 1377 (2010).
- Ross, S. M. Introduction to probability models. Access Online via Elsevier (2006).
- Scott, J., Gass, R., Crowcroft, J., Hui, P., Diot, C. & Chaintreau, A. CRAWDAD data set cambridge/haggle (v. 2006-09-15). Downloaded from <http://crawdad.cs.dartmouth.edu/cambridge/haggle> (2006).

Acknowledgements

This work was Supported by the Collaborative innovation project of "Integration of Moss, digital intelligence" funded by the innovation fund of University production, learning and research of the science and technology development center of the Ministry of Education. The Natural Science Foundation of Hubei Province under Grant No.2016CFC721, No. 2019CFC881 and 2019ZYYD009. The Young Talent Project of Hubei Provincial Department of Education under Grant No. Q20162801. The Project of Hubei University of Science and Technology under Grant No.ZJ0016, No.2019HX026 and No.BK202030.

Author contributions

H.C. wrote the main manuscript text and C.W. prepared the experiment and all the figures. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021