Inter-Session Network Coding with Clustering Routing in Wireless Delay Tolerant Networks

Yang Qin*, Weihong Yang, and Lei Peng
Department of Computer Science,
Harbin Institute of Technology (Shenzhen),
Shenzhen, China,
*corresponding author: csyqin@hit.edu.cn

Abstract—**Delay Tolerant Network (DTN) is a multi-hop wireless network, which has an intermittent connection due to the mobility of nodes, short of wireless communication range, sleeping mode of nodes. As a result, delay tolerant network usually gets higher average delay and overhead. In this paper, we propose an Inter-session Network Coding based Clustering routing (INCC). Inter-session network coding codes messages from different flow together using the broadcast nature of wireless network to improve the transmission efficiency. Due to the dynamic topology structure, the inter-session network coding scheme cannot be applied to DTNs directly. Therefore, we cluster the node to gain more coding opportunities. We classify the nodes into different clusters according to their contact frequency. The packets could be performed network coding in the proposed scheme even if they have different destination nodes. Then, we use Ordinary Differential Equations (ODEs) to investigate the advantage of applying network coding in DTNs in terms of average delay. Finally, we evaluate the performance of INCC via simulations, and simulation results show that INCC can achieve lower average delivery delay and reduce the network load.**

Keywords—network coding; buffer management; routing protocol; delay tolerant network

I. Introduction

Delay tolerant network is a multi-hop wireless network which has an intermittent connection due to the mobility of nodes, short of wireless communication range, and sleeping mode of nodes. As a result, delay tolerant network usually gets higher average network delay and overhead ratio. However, delay tolerant network could meet the requirements of some extreme environment, such as vehicle ad hoc network. Researchers try to improve the transmission efficiency of DTNs in two ways: the first one is to design an efficient routing protocol, and the second is to manage the cached data of the node efficiently. Network coding technology is very popular and it has a natural advantage of improving the transmission reliability, reducing the delay, improving the throughput, especially suitable for high packet loss rate of the wireless network.

In this paper, we propose an Inter-session Network Coding with Clustering routing (INCC) in DTNs according to the feature of inter-session network coding. Since inter-session network coding is a scheme which creates coding opportunities by detecting the topology of the network, it is difficult to be applied in DTNs directly. We classify the nodes into different clusters based on their probability of encounter to increase coding opportunities. Analytical results show that clustering routing can achieve higher delivery ratio than the non-clustering one. Simulations are performed to evaluate INCC by comparing with other algorithms.

The rest of the paper is organized as follows. Section II reviews the related work. Section III proposes the inter-session network coding with clustering routing and analyzes the performance of the proposed scheme. Section IV shows the simulation results and analysis. Section V concludes the paper.

II. Related Works

Routing strategies in DTN can be classified into two categories: stateful and stateless. In stateless routing, nodes do not have to keep routing information [1-3]. For example, Epidemic [1] is a flood-based routing and can achieve low delivery delay. However, it introduces redundancy and consumes resources, resulting in a heavy load and low delivery ratio when resources are limited. Spray and Wait [2] is a quota-based routing, which combines the diffusion speed of Epidemic and the simplicity of direct transmission to reduce the overhead of flooding-based schemes. However, the inefficiency of Wait phase can result in high average delivery delay. In stateful routing, nodes have to keep routing information [4-7]. For example, Prophet [4] and ProphetV2 [5] are routing protocols that forwards messages to neighbors based on their delivery probability to the destination. Abdelkader et al. [6] use a global mathematical model for optimal routing in DTN. However, the global information of the network may not be available.

In order to improve the efficiency of routing, researchers adapt network coding in DTN. Several proposals investigate the possibility of applying random linear network coding in DTN [8-15]. For example, Alia et al. [9] apply Random Linear Network Coding (RLNC) to improve the performance of reliable transport DTNs. Reference [11-12] investigate the performance of RLNC based Epidemic routing and show the significant performance improvement for data transmission in DTNs by using RLNC. MORE [14] applies intra-session network coding in opportunistic routing. Most of the existing works apply RLNC in DTNs, RLNC is an intra-session network coding, namely, only the packets that have the same source and destination can be performed network coding. COPE [15] is an architecture for wireless mesh networks, which uses inter-session network coding to increase the information content of each transmission and results in high network throughput and low delay.

III. Inter-session Network Coding with Clustering Routing

A. Design of inter-session network coding with clustering routing

In inter-session network coding, packets with the same destination are coded, then the coded packet is decoded on another node by using the broadcast nature of wireless network or receiving other coded packets. But DTNs is a spare network, packets usually have different destinations. So we cannot apply inter-session network coding to DTNs directly. Therefore, we introduce clustering into inter-session network coding based routing. Consider the topology in Fig. 1, the packets forwarded to the same destination (e.g., node *C*) can be coded on node *R*. Node *C* can decode the packet if it receives a native packet from node of *A* (i.e., P_1) or *B* (i.e., P_2), or any independently coded packet from another node.

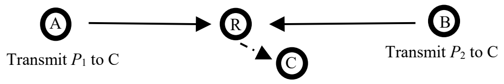


Fig. 1. Two packets destined to the same destination.

In Fig. 2, without clustering, packets get no coding chance at node *R*. After applying clustering, packets that are forwarded to the same cluster can be coded in a certain node like node *R* and then forward to nodes in cluster *C*. Any nodes in cluster *C* receive two linear independent coded packets can get P_1 and P_2 from the coded packet and forward the native packet to its real destination.

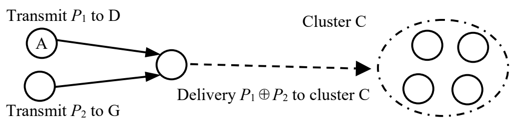


Fig. 2. Clustering routing in DTNs.

B. Analysis and design of clustering algorithm

1) Analysis of delivery ratio

We denote the set of nodes that storing message *m* as N_m, $\lambda_{i,d}\left(i \in N_m\right)$ denotes the probability that nodes with message *m* meet destination *d*, TTL_m is the lifetime of message *m*. The time interval U of a node carrying message *m* meets its destination follows the exponential distribution with parameter λ. Then, the probability that a node fails to delivery massage *m* to its destination within its lifetime is

$$P(U > TTL_m - t) = 1 - P(U \le TTL_m - t) \\ = 1 - \int_0^{TTL_m - t} \lambda_{i,d} e^{-u\lambda_{i,d}} du = e^{-\lambda_{i,d}(TTL_j - t)} \quad (1)$$

Then, the probability that message *m* can be successfully delivered is

$$P = 1 - P' = 1 - \prod_{i \in N_m} e^{-\lambda_{i,d}(TTL_m - t)} \quad (2)$$

Assume node *i* and node *j* are in the same cluster, node *i* and node *k* are in different clusters, then $\lambda_{ij} >> \lambda_{ik}$ ($>>$ denotes much greater than). Assume that node *i* is the source node of message *j* and node *d* is the destination. There are at most two replicas of message *m* in the network. We consider the case in which source *i* and destination *d* are in the same cluster. If the message *m* is only forwarded in the same cluster, then the probability that message *m* can be successfully delivered is

$$P_1 = 1 - e^{-(\lambda_{i,d} + \lambda_{j,d})(TTL_m - t)} \quad (3)$$

If the message *m* is forwarded in different clusters, then the probability that message *m* can be successfully delivered is

$$P_2 = 1 - e^{-(\lambda_{i,d} + \lambda_{k,d})(TTL_m - t)} \quad (4)$$

Since $\lambda_{j,d} >> \lambda_{k,d}$, we have $P_1 > P_2$. The probabilities of the case that the source *i* and the destination *d* are in different clusters can be derived similarly. We conclude that when the source and the destination of message belong to the same cluster, delivering the message in the same cluster can achieve higher successful delivery probability. Clustering can improve the probability of successful delivery. Moreover, clustering improves the performance of DTNs via increasing opportunities for nodes to perform inter-session network coding.

2) Design of clustering algorithm

Nodes are divided into different clusters and nodes leave or join the cluster based on probability of encounter. The probability between a pair of nodes is updated according to (5) and (6) as in [4]. We set α to 0.75 and γ to 0.98 in this paper.

$$p_{ij} = (p_{ij})_{old} + (1 - (p_{ij})_{old})\alpha \quad (5)$$

$$p_{ij} = (p_{ij})_{old}\gamma^k \quad (6)$$

The stability of the node *i* is defined as the average encounter probability between the node *i* and other nodes in the same cluster, which is calculated as follows

$$S_i = \sum_{j \in C_i} p_{ij} \Big/ N_i - 1 \quad (7)$$

The cluster updating is triggered by overtime and nodes encounter. If a pair of nodes does not encounter each other in a while, their encounter probability will be updated by (6). We denote the cluster of node *i* as C_i and the cluster of node *j* as C_j. When node *i* and *j* are in the same cluster, the node who has smaller stability value leaves the cluster. If nodes are in different clusters, we calculate the stability of node *i* in cluster C_j (denoted by S_{ij}) and the stability of node *j* in cluster C_i (denoted by S_{ji}), respectively. Node *j* will join C_i if $S_{ij} < \omega$; node *i* will join C_j if $S_{ji} < \omega$. The information about the updated cluster will be spread across the whole network.

C. Delay analysis of inter-session network coding with clustering routing

We use the Ordinary Differential Equations (ODEs) to analyze the impact of inter-session network coding on average delivery delay. Assume DTN network is composed of *N* nodes. The buffer size of each node is *M*. The probability of encounter is λ . T_i is the time that destination node receives packet *i*, $F_i(t)$ is the Cumulative Distribution Function (CDF) of T_i . V_i represent the set of nodes which has *i* messages in the buffer, $0 \le i \le M$. $B_i(t)$ denotes the number of nodes which has *i* messages in the buffer and at any moment we have

$$\sum_{i=0}^{M} B_i(t) = N \tag{8}$$

$R_i(t)$ denotes the rate at which a node in set V_i receives a new packet at the time *t* for coding DTNs routing scheme. The probability that a node receives a linearly independent coding packet is close to 1 when the coding domain is large enough [16]. $R_i(t)$ is calculated as

$$R_i(t) = \lambda(\sum_{j=1}^{M} B_j(t) + 1 - 1) = \lambda \sum_{j=1}^{M} B_j(t) \tag{9}$$

In a very short period of time *dt*, the probability of two nodes exchanging two packets or more is close to 0. Thus, for the non-coding routing scheme, we can derive the formula about the rate of $B_i(t)$ as follows.

$$dB_i(t)/dt = R'_{i-1}(t)B_{i-1}(t) - R'_i(t)B_i(t),\ 2 \le i < M \tag{10}$$

where $R'_i(t)$ denotes the rate at which a node in set V_i receives a new packet at *t* for non-coding DTNs routing scheme, and $R'_i(t) = \lambda[\sum_{j=1}^{M} B_j(t)\Pr(i,j) + 1 - \Pr(i,i)]$, $\sum_{j=1}^{M} B_j(t)\Pr(i,j)$ means that node in set V_i receives a new packet from the node in set V_j with probability of $\Pr(i,j)$. $1-\Pr(i,i)$ represents that node can always receive a new packet from the source node but cannot receive a new packet from itself. By applying inter-session network coding in DTNs routing scheme, the node selects the newly received packets $P1$ and cached packet $P2$ to code and get $P1 \oplus P2$, then the number of the cached data packet in this node did not change. We assume that coding opportunity probability in the network is *p*, then we can derive $B_i(t)$ of the coding based scheme as follows.

$$\begin{aligned} dB_i(t)/dt = {} & p_1 R_{i-1}(t)B_{i-1}(t) + (1-p_1)R'_{i-1}B_{i-1}(t)(1-p) \\ & - p_1 R_i(t)B_i(t) - (1-p_1)R'_i B_i(t)(1-p) \end{aligned} \tag{11}$$

where p_1 is the probability that a packet is a coded packet, which can be estimated as

$$p_1 = p(\sum_{i=1}^{M} B_i(t) + 1) \Big/ \sum_{i=1}^{M} B_i(t) i \tag{12}$$

We could get the relationship between $F_i(t)$ and $R_i(t)$

$$dF_i(t)/dt = R_{i-1}(t)(F_{i-1}(t) - F_i(t)) \tag{13}$$

The above formula is valid for both coding and non-coding routing scheme. The only difference is the receiving rate $R_i(t)$. The probability that a node can receive *i* message at time *t* is

$$S_i(t) = \Pr(T_i < t, T_{i+1} > t) = F_i(t) - F_{i+1}(t) \tag{14}$$

Now consider the destination node has received data packets, the intermediate node need to delete the already received packets. The scheme of deleting packet is Vaccine and Immune [17]. This paper only considers the Immune scheme in the non-coding routing scheme. In the coding scheme, we consider deleting of both coded packet and non-coded packet; the deleting of coded packet should happen when the destination receives all the coded packets. We consider the case in which $M = 2$, namely, each node can store at most 2 messages. We solve ODEs of the non-coded routing and coded routing scheme in Matlab with parameter $\lambda = 0.00003$ 、N=200, and the result is shown in Fig. 3. Due to the limited pages, we do not present the ODEs of the non-coded routing and coded routing scheme here. We observe the average packet delivery delay of the coding based routing scheme is lower than the non-coding based routing scheme.

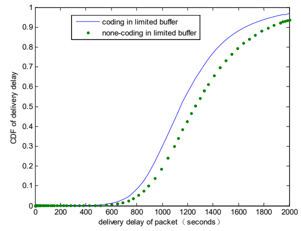


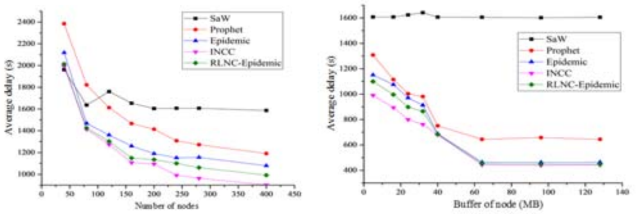
Fig. 3. CDF of coding scheme and non-coding scheme

IV. SIMULATION EVALUATION

We use the Opportunistic Network Environment simulator (ONE) in the simulation. We compare our proposal with Epidemic [1], Spray and Wait [2], Prophet [4] and RLNC-Epidemic [11] in terms of average delay, delivery ratio, and load balance. There are two types of nodes moving in a terrain of 4500m $\times$ 3400m. One type of nodes move with a velocity between 1m/s and 5m/s; the other move with a velocity between 5m/s and 10m/s. All nodes have a radio range of 10m and the bandwidth for transmitting data is 10Mbps. We use the Random Waypoint Model (RWP) [18]. The size of the packet varies from 50KB to 100KB and one new packet is created every 25 to 35 seconds. The simulation time is set to 43200s.

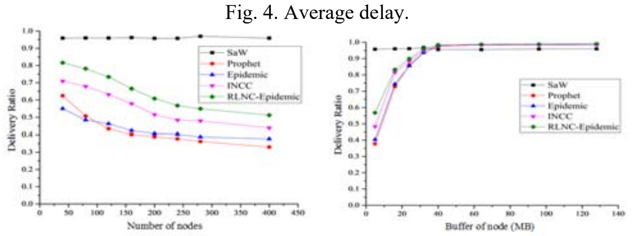
Fig. 4 shows the result of average delay under various numbers of nodes and various sizes of buffer. In Fig. 4(a), the buffer of each node is 5 MB. INCC can achieve the lowest delay among all the algorithms, which is consistent with analytical result in Section III. In Fig. 4(b), he number of the node is 240. We observe that INCC has the lowest delay. Moreover, the buffer size has minor impact on delay when size of buffer increases to a certain degree.

Fig. 5 shows the delivery ratio under various numbers of nodes and various sizes of buffer. In Fig. 5(a), the buffer size is 5MB. SaW has the highest delivery ratio because it limits the number of packet copies. INCC performs better than Prophet and Epidemic in that INCC increases the probability of successful delivery by clustering. In Fig. 5(b), the number of nodes is set to 240. SaW outperforms other four algorithms when the buffer size is small; however, when buffer increases, the advantage of SaW is not obvious. SaW can control the number of copies of message so that it has a good performance even the buffer is limited. When the buffer size of node becomes larger, it has little impact on delivery ratio.



(a) Different number of nodes (b) Different size of buffer

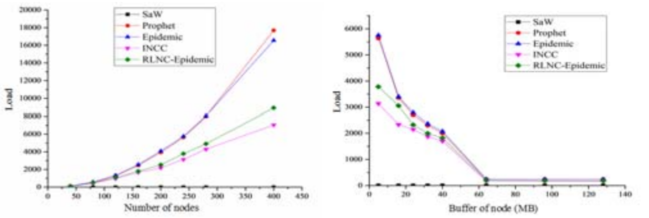
Fig. 4. Average delay.



(a) Different number of nodes (b) Different size of buffer

Fig. 5. Delivery ratio.

Fig. 6 shows the network load under various numbers of nodes and various sizes of buffer. The network load is defined as the ratio of the different between number of relayed packets and the number of delivered packets to the number of delivered packets. In Fig. 6(a), node's buffer is 5MB. Besides INCC, the load of other algorithms increases significantly with the incremental of node because more packets are generated, and the copies of packet will also increase. In Fig. 6(b), compared with other four algorithms, the advantage of INCC is not obvious in terms of load when the buffer becomes larger.



(a) Different number of nodes (b) Different size of buffer

Fig. 6. Delivery ratio.

V. Conclusion

In this paper, we propose INCC for DTNs. Since inter-session network coding creates coding opportunities by detecting the topology of the network, it is difficult to be applied in DTNs directly. Therefore, we divide nodes into different clusters to increase the opportunity for performing network coding. Then, we analyze the performance of clustering in terms of delivery ratio. We use ODEs to investigate the advantage of applying network coding in DTNs and analytical results show that the average delay of the inter-session network coding scheme in DTNs is decreased compared with the non-coding scheme. Finally, several simulations are performed to evaluate INCC by comparing with other algorithms. INCC performs well in terms of average delivery delay and can reduce the network load.

Acknowledgment

This work was supported by the Science and Technology Fundament Research Fund of Shenzhen under grant JCYJ20170307151807788.

References

[1] A. Vahdat, D. Becker et al., "Epidemic routing for partially connected ad hoc networks," 2000.

[2] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in Proceedings of the 2005 ACM SIGCOMM workshop on Delay tolerant networking. ACM, 2005, pp. 252–259.

[3] S.-C. Lo and C.-L. Lu, "A dynamic congestion control based routing for delay-tolerant networks," in Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on. IEEE, 2012, pp. 2047–2051.

[4] A. Lindgren, A. Doria, and O. Schel´en, "Probabilistic routing in intermittently connected networks," ACM SIGMOBILE mobile computing and communications review, vol. 7, no. 3, pp. 19–20, 2003.

[5] S. Grasic, E. Davies, A. Lindgren, and A. Doria, "The evolution of a dtn routing protocol-prophetv2," in Proceedings of the 6th ACM workshop on Challenged networks. ACM, 2011, pp. 27–30.

[6] T. Abdelkader, K. Naik, and A. Nayak, "Choosing the objective of optimal routing protocols in delay tolerant networks," in Computer Engineering Conference (ICENCO), 2010 International. IEEE, 2010, pp. 16–21.

[7] J. Burgess, B. Gallagher, D. D. Jensen, B. N. Levine et al., "Maxprop: Routing for vehicle-based disruption-tolerant networks." in Infocom, 2006.

[8] J. Widmer and J.-Y. Le Boudec, "Network coding for efficient communication in extreme networks," in Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking. ACM, 2005, pp. 284–291.

[9] A. Ali, M. Panda, T. Chahed, and E. Altman, "Improving the transport performance in delay tolerant networks by random linear network coding and global acknowledgments," Ad Hoc Networks, vol. 11, no. 8, pp. 2567–2587, 2013.

[10] X. Zhang, G. Neglia, J. Kurose, D. Towsley, and H. Wang, "Benefits of network coding for unicast application in disruption-tolerant networks," IEEE/ACM Transactions on Networking (TON), vol. 21, no. 5, pp. 1407-1420, 2013.

[11] S. Qin and G. Feng, "Performance modeling of network coding based epidemic routing in dtns," in Wireless Communications and Networking Conference (WCNC), 2013 IEEE. IEEE, 2013, pp. 2057–2062.

[12] Y. Lin, B. Liang, and B. Li, "Performance modeling of network coding in epidemic routing," in Proceedings of the 1st international MobiSys workshop on Mobile opportunistic networking. ACM, 2007, pp. 67–74.

[13] S. Ahmed and S. S. Kanhere, "Hubcode: hub-based forwarding using network coding in delay tolerant networks," Wireless Communications and Mobile Computing, vol. 13, no. 9, pp. 828–846, 2013.

[14] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, Trading structure for randomness in wireless opportunistic routing. ACM, 2007, vol. 37, no. 4.

[15] S. Katti, H. Rahul, W. Hu, D. Katabi, M. M´edard, and J. Crowcroft, "Xors in the air: Practical wireless network coding," in ACM SIGCOMM computer communication review, vol. 36, no. 4. ACM, 2006, pp. 243-254.

[16] Z. Liu, C. Wu, B. Li, and S. Zhao, "Uusee: Large-scale operational ondemand streaming with random network coding," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–9.

[17] Z. J. Haas and T. Small, "A new networking model for biological applications of ad hoc sensor networks," IEEE/ACM Transactions on Networking (TON), vol. 14, no. 1, pp. 27–40, 2006.

[18] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," Mobile computing, pp. 153–181, 1996.