A Named Data Approach for DTN Routing

Yaoxing Li[1,2] Yuhong Li[2] Lars Wolf[1] Anders Lingren[3,4] Ji Wang[2]
[1]Institute of Operating System and Computer Networks
Technische Universität Braunschweig, Braunschweig, Germany
[2]State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications, Beijing, China
[3]SICS Swedish ICT, Sweden
[4]Luleå University of Technology, Sweden

Abstract: Many common DTN routing protocols are replication-based, which have relatively good performance in terms of message delivery ratio but high overhead, and leave the issue of garbage collection open. In this paper, we propose Named Data Distance Routing (NDDR), a named data based DTN routing approach which makes routing decisions for named data based on topological distance information. This helps to reduce the overhead of routing. We have implemented NDDR in the ONE simulation environment and the simulation results show that the proposed routing method has better performance in terms of message delivery ratio and network overhead compared with several typical replication-based DTN routing protocols.

Key words: Delay-and-Disruption Tolerant Network (DTN), Named Data Network (NDN), Routing

I. INTRODUCTION

Delay-and-Disruption Tolerant Networking (DTN) [1] is an approach which addresses the technical issues in heterogeneous networks lacking continuous end-to-end connections among nodes, e.g., due to node mobility, constrained power sources, or unreliable links. With the increasing scope of DTN applications, such as vehicular networking, monitoring of wild animals [2] and roadside noise [3], efficient and reliable data forwarding is becoming more important and challenging. DTN routing approaches should not only consider the message delivery ratio, but also the network resource consumption, especially in scenarios with frequent and long duration periods of disconnection.

Many routing protocols for DTN have been developed [4] and the performance of those protocols has also been investigated [5]. Among them, Epidemic [6], PRoPHET [7] and Spray-and-Wait [8] are well-known and often referenced routing protocols for DTNs due to their distinguishing behavior. However, there are some common problems regarding these protocols. First, many popular DTN routing protocols are to some extent replication-based such that several copies of a message are kept in the network in order to improve the probability of successful message delivery. Yet, this method leads to an unnecessary waste of nodes' resources. Second, many replication-based approaches lack the capability of garbage collection, that is, although a message has been delivered successfully by a node, the buffer resources associated with replicas of that message in other nodes cannot be freed automatically even though these copies are not useful anymore. Moreover, such messages are forwarded continually, wasting even more resources.

Named Data Networking (NDN) [9] is an information-centric networking paradigm, considering data as the major object in its communication semantics. The information-centric approach is also applicable in many DTN scenarios such as VANETs, wildlife monitoring, etc., where focus is also on retrieval of data content rather than on specific physical hosts. This inspires us to design a DTN routing protocol using ideas from NDN. In addition, NDN's hop-by-hop forwarding mechanism based on named data is well suited to the intermittent connection environment of DTNs, and NDN's in-network caching mechanism can be used to minimize the number of message transmissions in the network. Moreover, the pull-based data retrieval mechanism can also be used to increase the efficiency and reliability of data delivery.

Meisel et al. [10] argue that NDN automatically embraces ad-hoc networking and delay-tolerant data delivery. Mobile nodes can communicate based on what data they need, instead of computing a specific path to reach a specific node. They proposed Listen First, Broadcast Later (LFBL) [11], a data delivery protocol designed for networking via named data in ad hoc networks. Tyson et al. [12] explore the potential of integrating information-centric and delay-tolerant principles into a shared ICDTN architecture. They argue that ICNs allow the network to gain a better understanding of the data itself, and the caching and replication of ICNs can offer information resilience, which offers a huge potential in disrupted environments. In order to study the quantitative benefits of integrating the two principles, they instantiated ICDTN in a human contact network. But ICDTN is more like a model of the merged architecture rather than a particular system with specific details.

In this paper, we propose a named data distance routing (NDDR) approach for DTN and study its performance in terms of message delivery ratio and overhead by comparing it with some popular DTN routing protocols. Theoretical analysis and simulation results show that the proposed DTN routing approach works well with respect to both message delivery efficiency and network resource consumption.

II. NAMED DATA APPROACH FOR DTN ROUTING

Several advantages can be achieved when applying named data to the design of DTN routing. First, as mentioned above, many DTNs are information-centric in nature. This accords with the motive of NDN, and the merits of NDN techniques, such as in-network caching can be integrated into DTNs. Second, identifying data by given names makes data reusable. In traditional replication-based host-centric DTN routing, data is identified by the endpoint identifier of its source and destination, so once it is successfully delivered, all copies remaining in the network are no longer useful. In contrast, if data are independent of source and destination, they can be reused by other nodes in the network. This could significantly reduce network overhead.

Finally, a receiver-based data retrieval mechanism is suitable to be realized in the highly dynamic environment typically faced by DTN with its inherent decoupling of sender and receiver over space and time. This extends current DTN routing protocols and can satisfy the needs of more DTN applications.

Based on these considerations, we propose NDDR (Named Data Distance Routing), which introduces the named data and pull-based message transmission model into DTN routing and makes forwarding decisions based on topological distance information.

NDDR routes packets based on data names and distance information. Unlike traditional DTN routing protocols, a message delivery in NDDR consists of two phases: request phase and response phase. During the request phase, a requester node puts the name of a required piece of data into an Interest packet (using the same terminology as in NDN) and sends it to the network via flooding. Once the Interest reaches a node that has the requested data, the response phase begins. The node returns the Data packet that contains the name and corresponding data content. On each hop of this phase, the forwarder broadcasts the Data to all the nodes in its transmission range. However, only the eligible nodes, as determined according to the algorithm below, will continue to forward the Data in order to reduce network overhead while keeping data delivery performance high.

Each NDDR router maintains three major components: a Distance Table (DT), a Content Storage (CS), and a Pending Interest Table (PIT). The DT is used to store the distance with other nodes in hops and consists of multiple distance entries. Each entry contains three values for each known active endpoint. For a router N and an endpoint E, the three values are: the identifier of E, the distance between N and E in hops, and the highest sequence number (seqnum) seen in a packet sent by E.

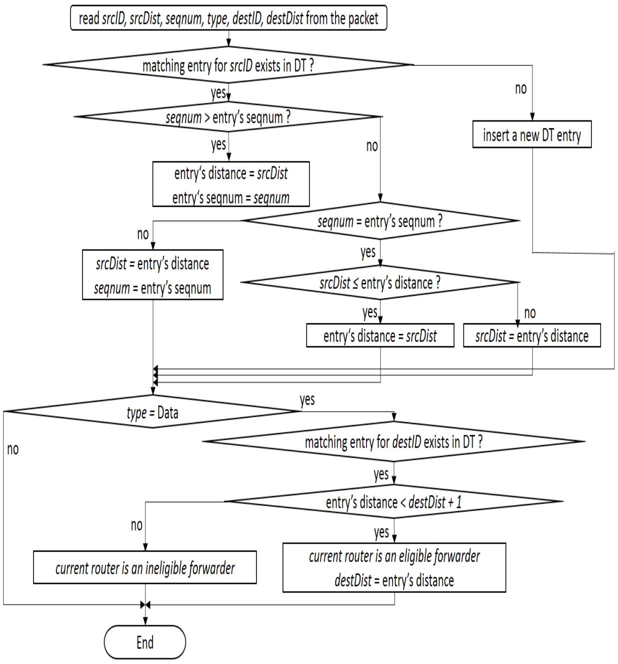


Fig. 1. Operations inside the distance table (DT) when a packet arrives

Fig. 1 describes the operations about how to update distance information inside DT when a packet arrives. PIT and CS serve the same purpose as in NDN. PIT stores all the Interests that have reached the router but have not been satisfied yet. CS caches all the Data that has reached the router for satisfying potential future requests.

When an Interest arrives, the NDDR router first checks its DT for a matching entry. If an entry exists, the router updates the distance information of the entry or the packet according to Fig.1. Otherwise, the router adds a new distance entry. Then the router checks CS for a matching entry by the data name. If the lookup hits, the router returns the matching Data and the processing of the Interest completes.

Otherwise, the router continues to use the data name to check the PIT. If there is a matching entry, the router adds the requester seen in the Interest to the entry's requester collection; otherwise it creates a new PIT entry, and increments the Interest's srcDist field, then forwards it. When a Data arrives, the router first updates the DT in the same way as it does when receiving an Interest. Subsequently, it checks CS; if a matching entry exists, it simply drops the Data, else it caches the Data. Finally, it checks the PIT and if a matching entry exists, the router replicates the Data and sends it to the requesters in the PIT entry's requester collection, and then removes the entry. Before forwarding, the router decrements each Data's destDist field and increases the srcDist field. If there is no matching PIT, the router drops the Data.

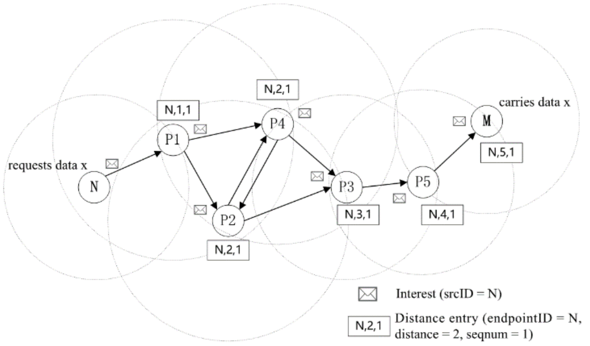


Fig. 2. Flooding during the request phase

Flooding during the request phase is shown in Fig. 2. Here, suppose node N requests a piece of data named x, and node M carries the data x. First, N creates an Interest: the dataName is x, srcID is N, and initial srcDist is equal to 0. Then N sends it to the network via flooding. Thus, the distance information of N is distributed - all the intermediate nodes along the way create a distance entry for N in their DTs to record the distance from themselves to N and an entry in their PITs to record that N's interested in data x.

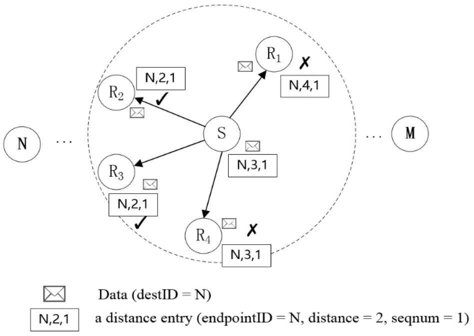


Fig. 3. One single hop of response phase

Fig. 3 depicts the forwarding process on one single hop during the response phase. Suppose sender S is forwarding the Data and receivers R_1, R_2, R_3 and R_4 are in S's transmission range. When they receive the Data from S, they will read the destDist which can be used to calculate the distance from S to N, namely $d_{S\text{->}N}$ = destDist + 1. Then they compare $d_{S\text{->}N}$ with the distance recorded in the N related entry in DTs. Obviously, R1 is farther to N than S, the distance from R4 to N is equal to that of S, while R2 and R3 are closer to N. R2 and R3 are eligible forwarders and they will forward the Data to the next hop. As ineligible forwarders, R1 and R4 will drop the Data.

III. IMPLEMENTATION AND EVALUATION

A. *Simulation Environment and Metrics*

We implemented NDDR in the ONE Simulator [13] which includes several common DTN routing protocols such as Epidemic, PRoPHET, and Spray-and-Wait. However, these protocols are push-based and host-centric, meaning that messages are sent from source to destination without requiring a return message (as needed for information request/response). To make the results more comparable, we extended the above three DTN routing protocols by simply adding a response message upon successful delivery of a request to ensure the round trip behavior.

We use the default map of the Helsinki downtown area (roads and pedestrian walkways) in our simulations. Table I shows the simulation parameters. We use default parameters for PRoPHET v1 and the value of L being 10 for Spray-and-Wait. We use SPMBM (Shortest Path Map-Based Movement) as the node movement model in the simulations.

TABLE 1 SIMULATION PARAMETERS

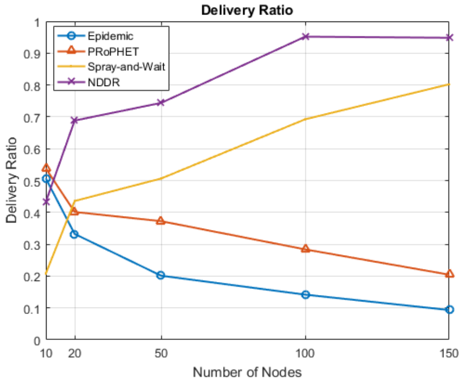
Parameter	Value
Routing approaches	Epidemic, PRoPHET, Spray-and-Wait, NDDR
Number of nodes	10,20, 50, 100, 150
Simulation duration	0.5 - 4 hours
Message creation interval	10 s
Message sizes	500kB - 1MB
Message TTL	Infinite
Node buffer size	50 MB
Node speed	13 - 15 m/s
Node transmission range	150 m
Mobility model	SPMBM

In each test when using Epidemic, PRoPHET and Spray-and-Wait, 10% of the nodes served as sources and another 10% served as destinations. After each message creation interval, a source is randomly selected from the source collection, and a message is created and sent to a randomly selected destination. For NDDR, 10% of the nodes served as requesters and another 10% served as responders.

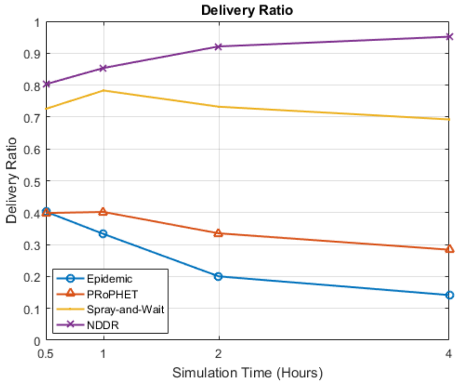
We evaluate the protocols using two metrics: message delivery ratio and overhead ratio. The message delivery ratio is the percentage of the original messages (Interest) that achieve the eventual delivery. For one original message (Interest), we consider the case that at least one corresponding ACK (Data) reaches the source (requester) as a successful delivery. The overhead ratio is the portion of transmissions used for something other than the successful delivery. It is calculated as the total number of transmissions of all the messages created during the simulation divided by the total number of transmissions of any message leading to eventual successful delivery, minus one.

B. *Evaluation on message delivery ratio*

Fig. 4 illustrates the delivery ratio of four protocols, where Fig. 4 (a) shows the effect of the number of nodes and Fig. 4 (b) shows the effect of simulation durations. In Fig. 4 (a), as the number of nodes increases, more copies are created when using Epidemic and PRoPHET. This leads to congestion and insufficient buffer capacity, thus the delivery ratio gradually goes down. As for Spray-and-Wait, the size of total buffer space occupied by copies is fixed. The more nodes there are, the larger the total buffer capacity is, and thus the larger probability that copies can stay in the nodes without being dropped until final delivery. For NDDR, the delivery ratio is only about 40% when the number of nodes is small. As the number increases, the delivery ratio gradually goes up and reaches a relatively high level of 90% when there are 150 nodes. NDDR makes forwarding decisions according to the distance information. When the network is sparse, the distance information carried by the nodes is unreliable since nodes may always encounter the others after a relatively long time and could not update distance information frequently. Consequently, as the number of nodes increases, the network is getting connected more and more tightly, the delivery ratio is improving.



(a) Delivery ratio vs. the number of nodes (simulation duration = 4 hours)



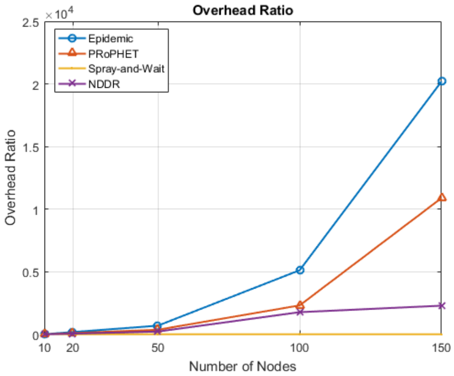
(b) Delivery ratio vs. simulation durations (number of nodes = 100)

Fig. 4 Delivery ratio of the NDDR

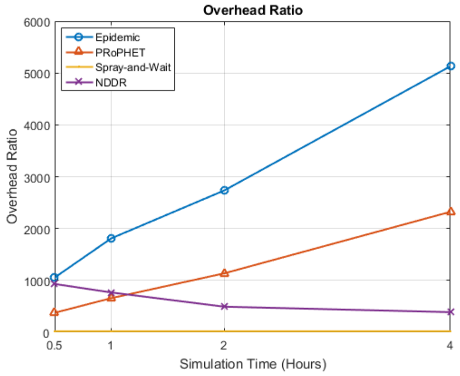
In Fig.4 (b), as the duration increases, the delivery ratio of Epidemic and PRoPHET is gradually decreasing. As longer the duration gets, as more opportunities for nodes to encounter others occur. When the number of original messages and copies becomes larger, heavier congestion and contention may happen, which results in the decrease of the delivery ratio. PRoPHET's message delivery is better than that of Epidemic since delivery predictability has been used, so it has a higher delivery ratio. The delivery ratio of Spray-and-Wait was affected little by the duration and is kept in a relatively high level. In contrast, the delivery ratio of NDDR is better than those of the other three protocols for all durations. The main reason is that when the duration is long enough, sufficient data are cached in the CS of each node. Therefore, the requesters are more likely to get the data that they are interested in from a closer node and, hence, the trip traversed by the data is shorter, leading to the decrease of dropped messages and an increase of the delivery ratio.

C. *Evaluation on overhead ratio*

Fig. 5 illustrates the overhead ratio of these four protocols. The effect of the number of nodes on the overhead ratio is explored in Fig.5 (a). As the number of nodes increases, the overhead ratio of Epidemic and PRoPHET increase as expected, because more nodes lead to more copies and, thus, more overhead. However, the number of nodes has no effect on the number of copies in Spray-and-Wait and buffer space is always sufficient for Spray-and-Wait. This is because whether the message will be successfully delivered or not is mainly depend on the last hop to the destination, namely, the only one hop of the wait phase, therefore, the result is always close to a constant, namely the number of copies. For NDDR, the caching in CS greatly shortens the distance messages have to travel since requests are always responded to by the closest responders. Moreover, PIT avoids unnecessary forwarding of requests (Interests) for the same data, which also contributes to the decrease of overhead. Furthermore, although NDDR creates multiple copies of messages, many of them do not traverse that far. As a result, the overhead is much less than those of Epidemic and PRoPHET for very dense networks.



(a) Overhead ratio vs. the number of nodes (simulation duration = 4 hours)



(b) Overhead ratio vs. simulation durations (number of nodes = 100)

Fig. 5 Overhead of NDDR

The overhead ratio of the four schemes at different simulation durations is illustrated in Fig.5 (b). As the duration increases, the overhead of Epidemic and PRoPHET increases. Spray-and-Wait creates a fixed number of copies. The buffer capacity is sufficient even when the duration is set to 0.5 hours, hence, the overhead ratio is kept at a low level. NDDR benefits from the in-network caching of NDN, and longer simulation runs gives more time for data to be located in the caches, leading to fewer hops to be travelled by the messages. Besides, as the simulation duration increases, there is a growing number of requests for the same data, i.e., PIT and CS are more effective, thus, the overhead ratio is gradually decreasing.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed NDDR, a named data based approach for DTN routing. Specifically, we introduced named data to the design of routing protocol and used the typical NDN communication model, i.e., the exchange of Interest and Data, to drive the communication, as well as the typical components of a NDN router, i.e., Pending Interest Table and Content Storage to serve as the functional components in NDDR routers. The simulations performed in the ONE simulator have shown that NDDR clearly gives better performance than Epidemic, PRoPHET and Spray-and-Wait with regard to message delivery ratio and overhead ratio. Overall, NDDR succeeds in its goal of simultaneously improving message delivery efficiency and reducing resources consumed during the delivery.

We have implemented a prototype of NDDR on top of IBR-DTN [14], a lightweight, modular and highly portable Bundle Protocol implementation and DTN daemon. In future work we will obtain more test results by establishing a test network as well as running more extensive simulations, including evaluating NDDR by using an advanced naming scheme, (e.g., hierarchically structured or flat names) for NDDR, and test it in a larger network. Furthermore, we want to study the implications of decoupling the NDN layer from the DTN routing layer so that any legacy DTN protocol can be made aware of cached data in the NDN layer and be used for forwarding of NDN data. This will enable the exploitation of the large body of research on DTN routing to be applied in different NDN scenarios without the need for new routing protocol development.

REFERENCES

[1] "Delay Tolerant Networking Research Group"; http://www.dtnrg.org

[2] P. Juang et al., "Energy-Efficient Computing for Wildlife Tracking: Design Trade-Offs and Early Experiences with Zebranet," SIGARCH Computer Architecture News, vol. 30, Oct. 2002, pp. 96–107.

[3] "Sensor Networking with Delay Tolerance (sendt)"; http://down.dsg.cs.tcd.ie/sendt

[4] S. Ali, J. Qadir, A. Baig,"Routing Protocols in Delay Tolerant Networks –A Survey", in Proc. of 6th International Conference on Emerging Technologies (ICET) , 2010, pp. 70 – 75.

[5] T. Abdelkader, K. Naik, A. Nayak, N. Goel and V. Srivastava, "A performance comparison of delay-tolerant network routing protocols, " in IEEE Network, vol. 30, no. 2, pp. 46-53, March-April 2016.

[6] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Technical Report CS-200006, Duke Univ. 2000. 6

[7] A. Lindgren, A. Doria, E. Davies, and S. Grasic, "Probabilistic Routing Protocol for Intermittently Connected Networks," Internet Research Task Force, RFC6693, August 2012.

[8] 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. 8

[9] L. Zhang, KC. Caida, P. Crowley, C. Papadopoulos, L.Wang and B.Zhang, "Named Data Networking", ACM SIGCOMM Computer Communication Review (CCR), July 2014.

[10] M. Meisel, V. Pappas, and L. Zhang, "Ad hoc networking via named data," in Proc. of the fifth ACM international workshop on Mobility in the evolving internet architecture. ACM, 2010, pp. 3–8. 3

[11] M. Meisel, V. Pappas, and L. Zhang, "Listen first, broadcast later: Topology-agnostic forwarding under high dynamics," in Annual conference of international technology alliance in network and information science, 2010. 9

[12] G. Tyson, J. Bigham, and E. Bodanese, "Towards an information-centric delay-tolerant network," in Computer Communications Workshops (INFOCOM Workshops), 2013, pp. 387–392. 4

[13] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE simulator for DTN protocol evaluation," in Proceedings of the 2nd international conference on simulation tools and techniques, 2009. 10.

[14] S. Schildt, J. Morgenroth, W.-B. Pöttner, and L. Wolf, "IBR-DTN: A lightweight, modular and highly portable bundle protocol implementation," in Electronic Communications of the EASST, Citeseer, 2011.