Prediction Based Routing With History Based Replication for Disruption Tolerant Network

K. Suresh and P. Ganesh Kumar
1Department of Computer Science and Engineering, 2Department of Information Technology, PSNA College of Engineering and Technology, Dindigul, India

Abstract: A delay and Disruption Tolerant Networking (DTN) is an emerging network, which handles communications in extreme environments, like space communications and networking in sparsely populated areas, vehicular adhoc networks and underwater sensor networking. DTN routing issues may appear as a standard dynamic routing problem with extended link failure times. To offset this, researchers investigated routing based on future contacts prediction, taking advantage of nodes’ mobility history. Routing is performed over time to achieve eventual delivery using long-term storage at intermediate nodes. In DTN, replication is used in DTN flooding strategies and there are many algorithms to manage multiple message copies and to make those copies. This study proposes a prediction based routing with history based simple replication for DTN routing.

Keywords: Delay and Disruption Tolerant Networking (DTN), flooding, history based replication, prediction based routing, replication

INTRODUCTION

DTN is a general-purpose overlay network operating over varying regional networks allowing regional networks with varied delay characteristics to interoperate through mechanisms to translate between respective network parameters. So, underlying protocols and technologies for regional networks differ considerably but DTN architecture’s flexibility allows them to be connected with each other. Presently, DTN architecture is based on the store, carry and forward paradigm (Bindra and Sangal, 2010). DTN supports regional networks interoperability by accommodating long delays between and in regional networks and translating between regional network communications characteristics. DTNs in providing such functions accommodate evolving wireless communication devices mobility and limited power (Kumar et al., 2012).

Various DTN features are followed by Fall and Farrell (2008):

- Intermittent connection
- High delay, high queue delay and low efficiency
- Limited resource
- Limited lifetime of node
- Dynamic topology
- Poor Security
- Heterogeneous interconnection

Deep space exploration, rural communication, studies of wild zebra, lake quality monitoring are some DTN applications (Wood et al., 2012). In September, 2003, Cisco router (CLEO) launched by satellite monitored the disaster in UK. Till December 2008, CLEO did routing tests in space environment using the Saratoga protocol of bundle layer instead of previous protocol, using the link source to overcome serious asymmetry link conditions. The Zebranet project installed a Global Positioning System (GPS) in a zebra collar to study zebra activities, an early DTN project started in 2004 (Zhang et al., 2004). There are rural communication projects in remote villages to provide access to Internet. Some try to reduce communications cost using asynchronous information transmission. European Union advises state and local governments to protect water quality activities where researchers didn't choose end-to-end communication mode, but using special node in the lake to cruise, realizing DTN storage and forwarding mechanism (McMahon and Farrell, 2009). There are other DTN application disciplines like industrial monitoring and disaster recovery (which are natural like volcanic eruptions, hurricanes, earthquakes or man-made disasters like terrorist attacks and car accidents). In all the above, wireless sensor nodes are simply required to periodically gather data related to the environment (heat, wind velocity, humidity, surface vibration, light intensity and noise) and report any major reading changes.
DTNs usefulness depends on network routing efficiency. DTN routing problems are constrained optimization problems where edges are unavailable for extended periods and each node has storage constraints. This formulation reveals DTN routing as different and more challenging (Chen and Murphy, 2001). Various DTN routing issues are routing objective, buffer space, resource allocation, reliability, energy and security.

Many classification schemes are proposed for DTN routing protocols. A common classification method is based on mobility behavior being deterministic or stochastic (Zhang, 2006). If nodes mobility behavior is deterministic and known, or at least predictable, then message transmission (where and when to forward packets) can be planned ahead of time to achieve optimal objectives. In a stochastic case, the network’s future topology (as a time-evolving graph) is assumed to be random or absolutely unpredictable. Thus, nodes are supposed to roam across a plane carrying data anticipating a suitable forwarding chance.

Generally, DTN routing protocols are classified into two categories based on which property finds the destination: flooding families and forwarding families. To find a destination, two different replication and knowledge approaches are used. Replication is used in flooding and there are many algorithms to manage multiple message copies and to make such copies (Jones and Ward, 2006).

Figure 1 shows a few routing examples, for both single and multiple message copies. An extreme case of multiple copies is flooding, where source and intermediary nodes deliver a message copy to every node which does not have the message. While single-copy algorithms need huge bandwidth for information collection, multiple-copy algorithms do so for message delivery (Islam and Waldvogel, 2008).

Flooding (random routing) is one of many stochastic approach variants based on available mobility information. When a node knows nothing about a network’s state, then it just randomly forwards packets to neighbor(s) (Jones et al., 2007). In flooding families, each node has copies of a message transmitting them to nodes set (also called relays). Relays maintain copies storing them in buffer space till they connect with next nodes. Direct contact, probabilistic routing, two-hop relay, epidemic routing, tree-based flooding, prioritized epidemic routing and Reconfigurable Ubiquitous Networked Embedded Systems (RUNES) routing protocols come under the flooding family (Vahdat and Becker, 2000; Ramanathan et al., 2007; Lindgren et al., 2003).

In forwarding families, network topology information selects best path and a message is forwarded from node to node on this path. Location-based routing, per-hop routing, source routing, per-contact routing and hierarchical routing protocols are from the forwarding family (Jones et al., 2007; Liu and Wu, 2007).

Wireless links are short-lived and end-to-end connectivity is sporadic in DTNs. Intermittent end-to-end paths and changing topology lead to failure of MANET routing protocols as they are designed under the assumption that network are connected. Nodes decide who the next hop is and when to forward, as the nodes route packets to destinations in a store-and-carry process. Earlier works fall into mobile resource-based, opportunity-based and prediction-based categories. This study proposes a prediction based routing with history based simple replication for DTN.

LITERATURE REVIEW

A survey on DTN, an architecture meant for Interplanetary Internet was presented by Puri and Singh (2013). DTN is a communication system providing Internet-like services in interplanetary distances to support deep space exploration. An introductory overview of Vehicular DTNs was provided by Pereira et al. (2012). This study addresses issues like routing and an introductory applications description and most important projects are given. Major DTN practical applications were reviewed by Wei et al. (2014) and focused on understanding social ties between nodes. Also, investigating design-related issues of social-based routing approaches like ways to obtain social relations among nodes, metrics and approaches to identify socialites characteristics, strategies to optimize social-aware routing protocols and suitable mobility traces to evaluate such protocols. DTN’s social properties in DTNs were summarized by Zhu et al. (2013), providing a survey of recent social-based DTN routing approaches. The methods take advantages of positive social characteristics like community and friendship to assist packet forwarding or consider negative social
characteristics like selfishness to improve routing performance.

Impact of delayed data transfer to end-to-end transfer and data life impact on network performance was examined by Sevimli and Soyturk (2010). Considering vehicles speed and desired throughput, data life is determined according to network connectivity. End-to-end delay bounds for data flows in DTNs were computed by Uddin et al. (2010). Results show that upper bound is moderate in pessimism and is good for deployment planning. Tradeoffs between packet delivery delay and various packet transportation types, a recurring theme in DTNs was shown by Tasiopoulos et al. (2014). Tradeoff between packet delivery delay and its transportation cost is captured on cost-delay plane using Optimal Cost/Delay Curve (OC/DC), for cases where a packet follows optimal routes and Achievable Cost/Delay Curve (AC/DC), where a specific (suboptimal) routing protocol is used. A metric called ExMin which stochastically calculates a metric by taking expectation of the minimum delays over all possible routes was proposed by Jeong et al. (2014). This study proved that online computation of ExMin was possible by relying on local information sharing.

A localized algorithm for DTN routing problems developed by Yang and Wu (2013) was inspired by the virtual backbone-based routing for mobile adhoc and sensor networks. In DTNs, which have intermittent connectivity, nodes meeting frequency was explored for backbone construction. It developed a Delay Tolerant Connected Dominating Set (DTCDS) as an approximation of DTN backbone and further formalized the problem of minimum equally effective DTCDS. A localized heuristic algorithm to build an efficient DTCDS is proposed. Studies include a theoretical analysis and comprehensive simulation of the new algorithm.

Road networks graph attributes was studied by Peng et al. (2013) by sampling real road networks in main Europe and USA cities. It also suggested a new graph metric, called characteristic central length, to estimate average shortest-path length of a large-scale spatial network. Huang et al. (2013) studied topology control in a predictable DTN, where time-evolving network topology is known a priori or can be predicted. This study models such time-evolving network as a directed space-time graph including spatial and temporal information. Topology control aims to build a sparse structure from original space-time graph so that:

- A network remains connected over time and supports DTN routing between any two nodes.
- Total structure cost is minimized, proving that this is a NP-hard and proposing two greedy-based methods that reduce total topology cost while maintaining connectivity over time.

A concept of adaptive routing, meaning choosing routing protocols based on DTN application characteristics was proposed by Liang et al. (2013). Adaptive routing schemes was developed according to specific requirements, history network information and current network conditions in network scenarios and validate that with adaptive routing. Network performances improve in many aspects like delay, overhead and delivery through simulations. Improving performance of adaptive routing from a resource allocation where bandwidth is critical and limited affecting routing performance was proposed by Liu and Wu (2009). It suggested an adaptive routing protocol, called Efficient Adaptive Routing (EAR), which allocates bandwidth (or forwarding opportunities) between multi-hop forwarding component and its mobility-assisted routing component dynamically to improve bandwidth usage. Simulations evaluated EAR routing performance under different network parameters.

A dynamic segmented network coding scheme to exploit the transmission opportunity efficiently when it is scarce in DTNs was proposed by Zeng et al. (2013). It adopts a dynamic segment size control making the segmentation adapt to network dynamics. A lower bound of anticipated delivery delay for bulk-data dissemination by using segmented network coding is also derived. An alternative Linear Programming (LP) approach for DTN routing was proposed by Amantea et al. (2013). This study showed that the formulation is equal to that presented in a seminar. But it has fewer LP constraints and a structure that suits Column Generation (CG) application. Simulation reveals that CG implementation provides an optimal solution of up to three orders of magnitude faster than original linear program in DTN examples.

Hajiaghajani et al. (2014) introduced a Gain-aware Dissemination Protocol (GDP) that attempts to reach maximum economic gain of content delivery through a balance between value achieved via message delivery and involved forwarding costs given out as user rebates. Economic gain from disseminating content is a generated value on content delivery minus forwarding costs. The key concept behind the new protocol is to balance adaptively between dissemination latency and forwarding costs to maximize economic gain. Using DTN simulation software ONE, a GDP protocol with varying mobility models, content generation times and content sources is characterized.

Correlation between a node’s meetings with other nodes was analyzed by Bulut et al. (2010) which focused on utilization of this correlation for efficient message routing. This study introduced a new metric called conditional inter-meeting time, that computed average inter-meeting time amid two nodes comparative to a meeting with a third node using local knowledge of past contacts alone. It also proposed...
routing messages over conditional shortest paths where link cost between nodes is defined by conditional intermeeting times. It suggested using conditional intermeeting time as an additional delivery metric when making message forwarding decisions. This trace-driven simulation on three datasets shows that modified algorithms are better than the originals.

Fair-Route, a DTN routing algorithm inspired by social processes of perceived interaction strength, where messages are forwarded to users with stronger social relation with message target was introduced by Pujol et al. (2009); assortativity, limiting messages exchange to users with similar "social status". This study compared Fair-Route performance to state-of-the-art algorithms through simulations on a MIT reality mining dataset. Results revealed that the algorithm outperformed current algorithms in a de facto benchmark of throughput vs. forwards. It distributes load better; top 10% carry out 26% of forwards and 28% handovers without performance loss.

A generic Prediction Assisted Single-copy Routing (PASR) scheme instantiated for different mobility models was proposed by Guo et al. (2013). PASR collects a short-duration trace with network connectivity information and uses an effective off-line greedy algorithm to characterize underlying network mobility patterns, depict best routing path features and guides on how to use historical information. It instantiates prediction assisted single-copy online routing protocols based on this guidance. So, instantiated protocols are energy efficient and are underlying mobility patterns aware. This study demonstrated PASR’s advantages in underwater sensor networks with varied mobility models.

A new, adaptive multi-step MDTN routing protocol was presented by Miao et al. (2014) where at every routing step, the protocol reasons on messages remaining time-to-live to allocate minimum copies to achieve a delivery probability. Results demonstrated the protocol as having higher delivery ratio and lower delivery cost compared to state-of-the-art Spray-and-Wait and Bubble protocols.

METHODOLOGY

Node mobility in prediction-based schemes is estimated on a history of observations. An example is utility-routing (Zhang et al., 2007), where a node has a utility value for every other node which is updated using time between contacts. A node forwards a message copy to nodes with higher utility for a message destination. Here source creates \( r \) identical message copies which are then delivered to “best” \( r \) nodes, where quality is history determined. Intermediate nodes then perform Direct Delivery.

History information predicts nodes’ future mobility in prediction-based routing schemes, which is the basis of decisions to forward messages to destinations. Earlier prediction-based DTN routing predicted whether two nodes would meet each other, but also consider when two nodes meet insufficiently. Two nodes meeting with a probability distribution improves delivery ratio and reduces delivery latency. PER is a single-copy DTN routing protocol-only one instance of a message is forwarded to the destination. Every message has in its header a Time-to-Live (TTL) field. When TTL expires, the message is dropped. Messages are forwarded hop-by-hop in succeeding contacts using a greedy approach. At every forward step, PER selects next hop with highest delivery probability to destination.

An additional metric studies various probabilistic delivery probability metrics \( f(x) \) during the packet forwarding. Distributions \( \Phi_{ij}^m(k) \) gives probability that future location at time \( k \) of a node \( m \) will be \( j \) considering that at time \( 0 \) location was landmark \( i \). Assuming that nodes trajectories are independent and that a recent known node state \( a \) is \( s_a \) (at time \( k_a \)) and for node \( b \) is \( s_b \) (at time \( k_b \), with \( k_a < k_b \), \( k_b < k \)), contact probability between \( a \) and \( b \) at a landmark \( i \) at time \( k \) is:

\[
C_{ab}(k) = \Phi_{ai}^a(k - k_a) \cdot \Phi_{bi}^b(k - k_b) \quad \text{for} \quad k > 0.
\]

Then, probability that \( a \) and \( b \) are in contact at a time \( k \) at a landmark is:

\[
C_{ab}(k) = \sum_{i \in L} C_{ab}^i(k) \quad \text{for} \quad k > 0.
\]

Note that \( C_{ab}(k) \) does not define proper probability mass function:

\[
0 \leq \sum_k C_{ab}(k) \leq 1
\]

For a probabilistic delivery study, probability metrics \( f(x) \), define probability that two nodes begin contact at time \( k \). When nodes \( a \) and \( b \) begin their first contact at time \( k \), it means that they had no contacts before an interval. A probability of first contact at time \( k \) assuming independent node trajectories is defined as:

\[
R_{ab}(k) = C_{ab}(k) \prod_{t=0}^{k-1} (1 - C_{ab}(t)) \quad \text{for} \quad k > 0.
\]

Denoting maximum message acceptance delivery delay by \( D \), meaning packets must reach destination in time \( D \).

History based simple replication (Jain et al., 2005) is used to predict nodes contact probability when
located. Reliability in protocols is by through acknowledgments and retransmissions. Timely feedback may be impossible due to DTNs intermittent nature, hence retransmission schemes have limited efficacy. An alternative is using replication and sending identical message copies simultaneously on many paths to mitigate single path failure (Jain et al., 2005). This is in contrast to retransmission schemes that wait for a message to be lost before forwarding a copy.

Identical copies are sent to first $r$ contacts as replicates, with $r$ being a replication factor. Only message source transmits or sends multiple copies, while relay nodes send or forward to destination alone; they cannot forward it to other relay mixing direct delivery and flooding. Source creates “$r$” identical message copies in history based simple replication and is delivered to “best” $r$ nodes, with history determining quality. Intermediate nodes carry out direct delivery (Jones and Ward, 2006) individually.

**RESULTS AND DISCUSSION**

The simulations are carried out for varying packet size (20-100), expiration TTL (20-120) and different buffer size (100-600). The delivery rate of the proposed prediction routing with history based simple replication is compared with prediction based routing.

It is observed from Fig. 2 that the delivery rate is consistently higher for the proposed prediction routing with history based simple replication when compared to prediction based routing. For expiration TTL of 20, the proposed method has 3.88% higher delivery rate than prediction based routing.

Figure 3 shows the delivery rate achieved for varying buffer size, it is seen that the proposed prediction routing with history based simple replication achieves higher delivery rate for buffer size of 200 to 500 when compared to prediction based routing. For buffer of 300, the proposed method has 8.57% higher delivery rate than prediction based routing whereas for buffer size of 100 and 600, the delivery rate is lower by an average of 1.29%.

It is observed from Fig. 4 that the delivery rate is higher for the proposed prediction routing with history based simple replication when compared to prediction based routing for different packet size. The proposed method achieves higher delivery rate than prediction based routing by an average of 3%.

**CONCLUSION**

This study proposes prediction based routing with history based replication; an efficient routing DTN scheme. DTNs allow network routing where contemporary end-to-end paths are either unstable or unlikely. Unstable paths are caused by link layer challenges. DTNs ensure messages though destination and/or intermediate nodes are offline. This delay is accepted for data exchanges. Protocols have varied knowledge levels for network nodes, from a set of nodes directly reachable over a node’s connectivity history and other nodes which predict the future accurately. Predicting nodes contact probability in the new method, history based simple replication is used when nodes are found. Simulations reveal higher delivery rate for new prediction routing with history based simple replication compared to prediction based routing for varying expiration TTL with a 3.65% on average. The new method achieved an average 3% higher delivery rate compared to prediction based routing for differing packet sizes.
REFERENCES


