Analysis of Community Behavior of Delay Tolerant Protocols

Ashish Kumar, Preeti Nagrath, Bharati VidyaPeeth

College of Engg.Delhi

Abstract

Delay tolerant networks constitute the category of Mobile Ad hoc Networks and are characterized by the ab-sence end-to-end path connectivity with limited data sources and power, where intermittent data communication is always a challenging task. To overcome this network partitioning, node mobility is exploited to increase message delivery in routing of these networks. Human mobility patterns have a great affect in increasing performance of routing protocol. In this paper, we have addressed, gathered and studied various routing protocols in DTNs that have used user mobility patterns for routing. These protocols use the constructive or destructive social characteristics for improving the performance in message forwarding. We have studied the impact of users social relationships on the protocols’ performance.

Index Terms

Community, Similarity, Friendship, asynchronous.

1. INTRODUCTION

ROUTING in delay tolerant environments is difficult as there is no guarantee that a fully connected path exists between nodes at any time and transfer opportunities are of limited duration. therefore conventional routing techniques applicable to MANET cannot be applied here. Different routing schemes have been proposed that are applicable to Delay Tolerant Networks. Many routing algorithms have been proposed to improve the routing performance. The main aim of a good forwarding strategy is to reduce the cost of forwarding the messages while retaining a high delivery rate. DTN have found to be useful in number of application such as: providing internet connectivity in remote areas where estab-lishing a complete infrastructure migh not be cost effective for example: DAKNet [26] Project, Wizzy digital courier service which provides asynchronous (disconnected) internet access to schools in remote villages. Wireless sensor networks [24] for wildlife tracking, a scenario where a hypothetical village is being served by digital courier service, a wired dial-up internet connection and a store and forward LEO satellite. Route selection through any one of them depends upon the variety of factors including message source and destination, size, time of request, available connects or other factors like cost and delay etc. [25] and transmission of information/ message during mission critical operations like natural disasters or battle zones. Due to its wide applicability it has become thrust area in network research. DTN posses several challenges due to limitation of resources such as bandwidth, buffer space, power supply, lack of infrastructure etc. The earlier strategies proposed for routing either rely on flooding based or knowledge-based techniques. Flooding based approaches works on the principle of replicating the messages and forwarding it to maximum number of nodes hereby increasing the chances of message delivery. Flooding the message in the network seems to be the most trivial approach. Attempts to offset the limitations of flooding based approaches like epidemic protocol[2] have introduced a host of other schemes. That is, a node copies its message to all the nodes that come in contact with it, provided the recipient node does not have a copy of it already. Epidemic is the simplest form of routing approach where in the replicated message is forwarded to every node coming in contact. It consumes network resources and is therefore not scalable. It gives good message delivery ratio with much less delay but is expensive in terms of resources since the network is essentially flooded. The epidemic approach has been improved by introducing controlled flooding approaches. For controlling the flooding in epidemic we have a protocol named Spray and Wait algorithm [4] that controls the flooding overhead by limiting the number of message copies flooded to the network thereby reducing the resource utilization. Some prediction based algorithms are also given to improve message delivery and reduce PROPHET[9] uses the history of encounters and transitivity to calculate the probability that a node can deliver a message to a particular destination A message is forwarded to a node if it has higher delivery predictability than the current node for that particular destination. JAIN et al [11], in their work proposed several routing algorithms, based on knowledge oracles. They proposed the Minimum Expected Delay (MED) protocol, which is based on the future contact schedule. Jones et al [3] improved this work by proposing Minimum Estimated Expected Delay (MEED), where the expected delay is computed using the observed contact history. Using a sliding window, each node records the connection and disconnection time of every contact. In this way, the most recent information is available for routing purpose, while MED has only one information. As humans always behave in a certain
specific mobility pattern so the concept of social based forwarding came. The DTN nodes may show social behavior and their movement might not always be random. Some of the approaches make use of the social behavior of the nodes. The nodes having common interest may meet frequently and may be more useful in transferring the messages. Due to the social behaviour the group based mobility models are becoming increasingly popular in DTN research area. The routing approaches exploiting social grouping performs better than various other approaches in terms of message delivery ratio, message traffic and message delay. In some of the applications it is desired to forward the message only through members of certain set of groups. Geographically distributed groups having common interest may form a community. It is desired to forward a community based message only through the nodes that are the members of groups of the community Bubble Rap protocol[4] is based on the popularity index(Rank) of the nodes. Each node carries two types of Ranking(Centrality) in the network: Global a global ranking (i.e.global centrality) across the whole system, and also a local ranking within its local community. If a node in community wants to send message to another node in other community then the source node first bubbles the message up the hierarchical ranking tree using the global ranking, until it reaches a node which is in the same community as the destination node, then the local ranking system is used instead of the global ranking, and the message continues to bubble up through the local ranking tree until the destination is reached. This protocol represents the static structure of the networks. Social awareness [1] goes beyond personal contexts and extends to group and community levels. The objective is to reveal social interaction (e.g., group detection, friendship prediction, situation reasoning) patterns, human mobility patterns, social situation (in a meeting, at a restaurant, meeting friends), and so on. In real social network some nodes are highly connected, some less and some least. Such a behavior can be used to improve the forwarding strategy to route messages to the destination.

2. PROPERTIES AND CHARACTERISTICS OF DTN NETWORK

In this section, we will discuss the various social properties and characteristics which specify the performance of DTN networks.

A. Centrality

Centrality is the measure of importance of the node within the network. A central node, typically, has a stronger capability of connecting other network members. A node with high degree centrality maintains contacts with numerous other network nodes. Such nodes can be seen as popular nodes with large numbers of links to others. It is the measure of the contribution of network position to the importance, influence, prominence of a node in a network. It helps in the mapping and measuring of relationships and flows between people, groups, organizations and computers. There are several ways to measure centrality, degree Centrality, betweenness centrality and closeness centrality. Degree Centrality is based on the simple calculation of how many links (or relationships) a node/member has. The greater the vertices or connections, the greater the Degree Centrality of that node. Betweenness Centrality indicates that a node or member appears often in the paths between other members. These are the connector types within the network, making introductions and maintaining many relationships between those who might not have otherwise connected. They are the glue that ties your community together. Closeness Centrality measures the physical distance from the node/member to the center of the network. Nodes that have high closeness may have a lot of traffic through them, but they're somewhat less valuable since they're easily replaced by other nodes/members with high closeness. In general, these are your core members or people within the network, they're connected to a lot of the other core members and they may have some connection to outlying groups but are rarely the sole path to those groups.

B. Popularity

Popularity defines the social feature of a node. Popularity of a node better known as, the node is interacting with different nodes each day. The higher a node’s popularity, the higher the chances of it meet more devices. Nodes with high betweenness and degree centrality measures are seen as nodes with high popularity. The relative popularity of a node is based on the number of connections and its ability to bridge the partitioned network. The concept of popularity is known as static popularity and dynamic popularity. Static popularity describes the connectivity of nodes in a predefined social network at the virtual level. Dynamic popularity refers to the social structure inferred from the observation of physical links over time. Differences may exist between static and dynamic popularity; thus impacting the identification of highly connected (or isolated) nodes. The characterization of the social network influences routing protocol performance.
C. Node Movement

Node Movement can be Static or Dynamic. Static mode provides fixed topology which doesn’t change inspite of node movement whereas Dynamic node movement allows the topology to change. Static nodes are Stationary or rarely move. Dynamic nodes moves frequently and hence the performance of the network changes with respect to the node movement. Dynamic node movement is more real as compared to static because all the devices which are used are considered to be rapidly moving.

D. Similarity

The congregation of mobile agents with similar characteristic patterns develops mobile societies in wireless networks. Researchers have long been working to infer these characteristics and ways to measure them. People with similar behavioral principle can be tied together. This brings an important aspect where, user-location coupling can be used to identify similarity patterns in mobile users. Similarity is an important aspect in DTN to develop behavioral space for efficient message dissemination and design behavior-aware trust advisors among others. For efficient networking, it can help to quantify traffic patterns and develop new protocols and application to target social networking. The degree of contact between nodes [6] has an important effect in terms of information dissemination. When the neighbours of nodes are unlikely to be in contact with each other, diffusion can be expected to take longer than when the degree of separation is lower. Analysis of similarity can be used to evaluate the network transitivity, which helps to analyze macro-mobility, evolutionary characteristics and emergent properties.

E. Friendship

Friendship is a social metric to define the social behavior of nodes. Friendship of a node with other nodes is either direct or indirect. Nodes can be considered friends of each other, if two nodes contact frequently, their contacts must be longlasting and regular. Even if two nodes are not meeting frequently but they are meeting once in a week (i.e. regularly) are also considered to be as friends. Two nodes can be close friends or weak friends. This behaviour feature of friendship is determined by frequency (how often they meet), longevity (how long they communicate with each other) and regularity (how regularly nodes meet).

F. Community

Community is well defined and very significant in Sociology and Biology. In Sociology, Community is a group in which the individuals who make up that group are motivated to take part in the group purely by self-interest. In biology, a community is a group of interacting living organisms sharing a populated environment. Now this concept is well adopted in wireless networks. Here nodes are considered as people and nodes of common interest are being clubed together to make a community.

![Community in Social Networks](image)

This figure shows how member of communities help in forwarding the data packets to the destination node D from Source node S.

G. Selfishness

Socialists and Economists have well defined the Selfishness in human behavior, which is well depicted and adopted in net-working. In Computer Networks (DTN), Selfishness attributed to the different behavior of a node to maximize its profit or to the node with which it has strong social ties. Social selfishness will affect node behaviors. As a forwarding service provider [23], a node will not forward packets received from those with whom it has no social ties, and it gives preference to packets received from nodes with stronger ties when the resource is limited. A node can be selfish at individual level or socially.

H. Data sets

The data set includes initial topologies of agents social networks and specication of knowledge networks for each of the agents to an empirically derived distribution of knowledge. Another task is creation of realistic task structures that could be used to simulate performance of complex interdependent projects by groups of agents. The main concern in generation of articial datasets is its realism. Social network datasets are extremely difficult to obtain and limited in size and scope. Social media datasets contain a large amount of
private information that can identify many facets of user identity and behavior. One of the premier reality data set are being provided by MIT. The Reality Mining project [28] represents the largest mobile phone experiment ever attempted in academia. They have collected an unprecedented amount of data on human behavior and group interactions that they plan on anonymizing and making available to the general academic community for further research.

3. SOCIAL-BASED DTN ROUTING

This section includes survey of many social and community based protocols.

A. Label

Hui and Crowcroft [11] had proposed a label routing method based on community labels in Pocket Switched network (PSNs). A PSN [12] is a subclass of Delay Tolerant Network and represents one particular intermittent communication standard for mobile radio devices. In Label routing Labelling strategy is used as Forwarding Scheme. It is considered that each node has a label telling others about its affiliation/group, just like the name badge in the conference. This scheme is entirely for forwarding messages to destinations, or to next-hop nodes belonging to the same group (same label) as the destinations. Hui postulates in this routing scheme that people from the same affiliation tends to meet more often than people outside the affiliation and hence can be good forwarders to relay messages to the other members in the same affiliation with the same label. This scheme required very modest information about each individual and is assumed to be very easy to implement in the real life, by just tapping a mobile device and write down the affiliation of the owner.

B. SimBet

To select a node which can be the best carrier for the message is a multiple attribute decision problem, where we have to select the node that provides the maximum utility for carrying the message. In SimBet[6] This is achieved using the similarity utility SimUtiln and the betweenness utility BetUtiln of node and compared with the other node for delivering a message to destination node. This protocol basically exploits the exchange of pre-estimated betweenness centrality metrics and locally determined social similarity to the destination node. Simulations using real trace data to demonstrate that SimBet Routing results in delivery performance close to Epidemic Routing but with significantly reduced overhead. Additionally, we show that SimBet Routing outperforms PRoPHET Routing, particularly when the sending and receiving nodes have low connectivity.

C. Dynamic Social Grouping (DSG)

Cabaniss et al. [7] proposed DSG algorithm which is strictly for node to base station. Its purpose is to identify a route to deliver messages to a base station. Delivery [7] must minimize the number of message repetitions while ensuring a high percentage of messages delivered to the destination. This algorithm uses social groups to improve communication throughput. Contact strength between nodes, Nodes probability of Delivery, Groups Probability of Delivery and Maximum Probability of Deliver are the variables which are considered to find the Individual message probability. DSG algorithm [7] has been shown to provide a significant increase in efficiency over probabilistic routing and epidemic routing. DSG has managed to achieve a high message delivery ratio and a low delivery time by reducing overheads. This will lead to better data aggregation and longer battery life.

D. Dynamic Social Grouping: Node to Node (DSG-N2)

Cabaniss et al. proposed [8] an algorithm known as DSG-N2. This algorithm is based on [7] and expanded to include node to node capability and improve grouping and routing efficiency. DSG-N2 performs better than Epidemic and SimBet routing in terms of total message traffic, delivery time and power consumption. The network recognizes cohesive social patterns and identifies them as groups, at the same time distributing this information throughout the network. DSG-N2 improves upon the probabilistic model by incorporating group probabilities. It uses minimum communications to reliably deliver the message. This algorithm emphasizes on minimizing the number of message repetitions, ensuring high delivery ratio.

E. Modified Distributed Bubble Rap Algorithm (DiBuBB)

Bubble rap algorithm defines two important aspects of social network centrality and community. BUBBLE [4] combines the knowledge of community structure with the knowledge of node centrality to make forwarding decisions. In Distributed Bubble Rap, device should be able to detect its own community and calculate its centrality values. It is discovered that the communities detected [15] by the distributed algorithms can satisfactorily approximate the centralised algorithms
which require the whole network topology. Among various other distributed algorithm k-Clique offers best performance. If the mobile devices can afford the storage for a local copy of the familiar set of its community members [15], k-Clique would be a good choice for its reasonably good similarity values and also quite low computational complexity, O(n²) in the worst case. DiBuBB uses distributed k-Clique [15] for community detection and a cumulative or single window algorithm for distributed centrality computation. The base version of DiBuBB is modified [14] to use the social network matrix instead of k-Clique. Thus, when two nodes meet,

$$4$$

instead of checking if they belong in the same community according to k-Clique, they look for a social link between them. If that social link exists, then the nodes will compare their community centralities, and the one with the lower value will send its messages to the other one. If there is no link between the nodes, the global centralities will be verified. This will be referred to as the social version. The centrality value is computed according to the following formula [14]:

$$\text{centrality} = w_1 \cdot S_{\text{window}} + w_2 \cdot \text{popularity},$$

where $S_{\text{window}}$ is the original value of the centrality as computed by DiBuBB, popularity is the number of social links a node has, and $w_1$ and $w_2$ are weight values that follow the conditions $w_1+w_2=1$ and $w_1 > w_2$.

F. Community Based Protocol

The protocol by Ma Xuebin[27] which is based on the community structure of node. The opportunity of one node meeting other nodes is not equal in Delay tolerant networks rather some nodes meet more frequently than others so we can make a community structure based on similarity criteria. The network can be seen as constituted by a number of connected communities with sparse links. Each node is given a community set based on the modularity and increase and decrease in this modularity decides whether a node joins a new network or still in the previous one. The two nodes on meeting increases a probabilistic metric and the node with maximum weight is used to forward the message. For simulation purposes one simulator is used. In order to increase the reality of node movement, the Working Day Movement (WDM) model is taken for the simulations. Simulation results show that this proposed protocol has higher delivery ratio and lower latency than Epidemic and PROPHET.

G. Social-stratification Probabilistic Routing Algorithm (SPRA)

Alnajjar and Saadawi [19] proposed an algorithm with the use of probabilistic routing in DTN architecture using the concept of social stratification network. The SPRA uses factor of social stratification of each node for forwarding strategy to solve the problem of periodically disconnected network. The procedure of SPRA relies on the knowledge of the mobility of nodes to forward messages based on social-stratification probabilistic procedure. The social stratification probabilistic factor of any node based on how many nodes did this node encounter until the moment of meeting a new node.

### Table 1

<table>
<thead>
<tr>
<th>Routing Protocols</th>
<th>Centrality</th>
<th>Popularity</th>
<th>Node Movement</th>
<th>Similarity</th>
<th>Neighborhood</th>
<th>Copies the intermediate nodes</th>
<th>Community centralities</th>
<th>Selfishness</th>
<th>Data mobility known as dispensed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lathe [1]</td>
<td>*</td>
<td>*</td>
<td>Static</td>
<td>*</td>
<td>Single</td>
<td>MIT reality mining datasets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SimBet [6]</td>
<td>*</td>
<td>Static</td>
<td>*</td>
<td>*</td>
<td>Single</td>
<td>MIT reality mining datasets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSG* [7]</td>
<td>Dynamic</td>
<td></td>
<td></td>
<td></td>
<td>Multiple</td>
<td>MIT reality mining datasets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiBuBB [14]</td>
<td>*</td>
<td>*</td>
<td>Dynamic</td>
<td>*</td>
<td>Multiple</td>
<td>Working day movement model data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community based Protocol [27]</td>
<td>Dynamic</td>
<td>*</td>
<td>Single</td>
<td>*</td>
<td>*</td>
<td>Real time data taken at Politecnica University</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPRA [19]</td>
<td>Dynamic</td>
<td>Single</td>
<td>Two Copies</td>
<td>*</td>
<td>Real time data</td>
<td>Real time data taken at Politecnica University</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GizaGet [20]</td>
<td>Dynamic</td>
<td></td>
<td>Two Copies</td>
<td>*</td>
<td>Infocom 05 [21], Cambridge 06 [22]</td>
<td>Infocom 05 [21], Cambridge 06 [22]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMART [17]</td>
<td>Dynamic</td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td>Vehicle DTN data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If node a meets node b and the social stratification probabilistic factor of node a is greater than node b, so it means that node a encountered more nodes than node b until the encountering time. In this case, nodes a will not forward any messages to node b. a node will forward messages to node b only if the probability of the encountered node is greater than the node that carried messages. SPRA uses the history of encountered nodes to predict its future suitability to deliver messages to next node toward destination. When two nodes meet, they update the summary vector. Then, they exchange summary vectors which in this case also contains the list of encountered nodes stored at the nodes. This information in the summary. Then, they exchange summary vectors which in this case also contains the list of encountered nodes stored at the nodes. This information in the summary. vector is used to decide which messages to request from the other node based on the social stratification factor used in the forwarding strategy.

H. GiveToGet Protocol (G2G)

Mei and Stefa [20] have taken into consideration the real environment scenario in which the nodes are selfish in nature. They have given two protocols Give2Get Epidemic Forwarding and Give2Get Delegation Forwarding. Both protocols are Nash equilibria, that is, no individual has an interest to deviate. G2G Epidemic Forwarding consists of three phases: Message generation, relay, and test. Message generation executes when one node has a message to send to some other node in the system. In relay phase precise design helps hide the sender of the message to every possible relay except the destination. This mechanism has the goal of making it impossible to another node to know whether it is the destination of the message or not before taking the message and giving the proof of relay. Therefore, In G2G Epidemic Forwarding, another node will follow all the relay protocol with the hope of being the actual destination of the message. In G2G Delegation Forwarding the proof of relay contains much more information, including the forwarding quality towards Destination claimed by another node and the forwarding quality of the message at that point in time.

I. A Secure Multilayer Credit-Based Incentive Scheme for Delay Tolerant Networks (SMART)

Zhu et al. [17] proposed SMART which is based on the concept of a naive scheme (layered coin) that provides virtual electronic credits to charge for and reward the provision of data forwarding in DTNs. The design goal of SMART is to provide a scheme which provides Effectiveness, Security, Efficiency, Generality. As the Message (bundles) forwards from Source, the bundle sender will introduce a base layer of the layered coin and also lose credits to the digital signature, to track the data forwarding path. It requires that the intermediate nodes opportunistically submit layered coins for clearance. The risk of the submission refusal attack is reduced by SMART by using multicopy forwarding. In charging and rewarding phase, source will decide the rewarding policy. The rewarding policy is propagated together with layered coin to each intermediate node. The charging model which is adopted in it is pay per packet. The computation and Transmission efficiency of smart scheme is improved by reducing the Transmission and Computation Overhead with an Aggregate Signature and efficient fragmentation authentication with the Merkle Hash tree. The successful delivery rate [17] of SMART is slightly lower in the beginning due to the extra security overheads, the network throughput would remain relatively stable since SMART can successfully stimulate selfish nodes in packet forwarding. It also proposed two efficiency optimization methods to reduce transmission and computation overhead.

4. CONCLUSION

Mobile devices are usually carried by humans, so the movement of such devices is necessarily based on human decisions and socialization behaviour. So it is depicted that user mobility helps routing and makes forwarding more efficient, that is why it has become important to model the behaviour of individuals moving in groups and between groups, as clustering, social grouping is likely in the typical ad hoc networking deployment scenarios. Various Features like Community, Friendship, Similarity, Centrality are constructive and help in forwarding messages whereas selfishness which is a destructive behavior makes routing inefficient. So in order to capture this type of behaviour, various protocols have been proposed that are heavily dependent on the structure of these social relationships among the people carrying the devices.

REFERENCES


