Social Geo-location Information (SGI) based Routing Protocol for Delay Tolerant Networks

Seungbae Kim, Ziyun Ye
Computer Science Department, UCLA

ABSTRACT

People enjoy sharing interesting contents or their daily life with location information through online social network services. Since this location information can be shared with friends, people know where their friends recently were. In this paper, we propose a routing protocol that uses location information from online social network services. Our work consists of two parts: It firstly downloads social information from Instagram then builds mobility pattern based on it. Every node has its own mobility pattern and knows mobility patterns of its friends. By comparing its own mobility pattern and that of destination node, we can calculate the probability that each node meets with destination and forward the message to those with higher probability. The simulation results show that the proposed routing protocol has similar delivery ratio but has 80% less network overhead compared with Epidemic routing protocol.

1. INTRODUCTION

In the last decade, many types of new Internet services have been emerged such as online social network services and mobile cloud services. The number of mobile devices explosively increases and it triggers the emergence of new types of applications as well. Almost everyone carries a smartphone and uses this device all day to talk to friends, share snapshots with others, or search something in Internet. Among these new types of services, online social network service is one of the most popular services. People can simply share their daily life, thinking, or feeling by posting a snapshot, video, or short texts with their friends. When people post something on their online social network, they can also provide location information of a specific photo or a place that the user is located in. This location information can be seen by their friends so that people can easily figure out the places that their friends have been by checking the name of the places on each posting. In this paper, we define this new type of information as ‘Social Geo-location Information’ (SGI). **Fig 1** shows an example of SGI in an Instagram photo.

In this paper, we present an efficient routing protocol for the delay tolerant networks (DTN) which is based on the social geo-location information. Many researches have been conducted to propose routing protocols for DTN but this is the first work that uses real location data from online social network services. Some previous researches [6,7] have proposed routing protocols that use social information. However, the sociality in these researches are more like an encounter probability or node degree in a social network. Additionally, there is an important factor that the previous DTN researches missed. In any networks, a source node and a destination node are not randomly picked, instead they talk to each other to request some data or send specific information. In previous researches, the traffic was generated from one random node to the one random destination. In real life however, people do not talk with strangers but their friends or people who already know. In this work, we define relationships among nodes so that only two nodes which have a friend relationship with each other can send/receive a message. Also we assume that people repetitively visit certain places and contact with their friends during daily activities. So everyone already knows their friends, and also their mobility patterns since online social network services kindly provide the location information for their friends.

Fig 1. Social Geo-location Information (SGI) in an Instagram photo

Our work have two contributions: First, we developed a SGI data collector which downloads friends' relationship information from online social network services and builds mobility patterns of the collected users. We chose the Instagram as our data source because people upload snapshots of their daily life with location information. We firstly collected photos that are taken at UCLA and its metadata, and get further friends user IDs and their recent photos. Photos that contain location data which is latitude and longitude allow us to make a mobility pattern of the user who uploaded the photos. Second, we proposed the SGI based DTN routing algorithm and evaluate the performance in terms of the delivery ratio and the network overhead. Since every node sends a packet to a destination node which is a friend and it already knows the mobility pattern of the destination node, the only thing that sender needs to do is to copy and forward the packet with the mobility pattern of the destination node. Whenever two nodes are encountered, the sender gives the mobility pattern of the destination node to the intermediate node, and the intermediate node calculates the chance to meet the destination by comparing its own mobility pattern with the destination’s mobility pattern. If the calculated probability is higher than that of source node, sender forwards the packet to the intermediate node. In this approach, intermediate nodes do not need to keep updating a large size table that includes the encounter probabilities or node degree of all nodes in a network. So that the process time on each node can be significantly decreased. Also our simulation results show that the proposed routing protocol has similar delivery ratio but has 80% less network overhead compared with Epidemic routing protocol.

This paper is organized as follows. In section 2, we introduce other DTN routing protocols that are used in our simulation and also protocols which use social information. In section 3, we explain how we collect SGI data from Instagram and how SGI routing protocol works. Then we evaluate the performance of our proposal by using simulation in section 4. Finally, we conclude the paper and discuss the future works in section 5.

2. RELATED WORKS

In this section, we present a short review of several efficient routing protocols for DTNs. One set of the algorithms is based on epidemic routing protocols [1], where nodes simply forward message to encountered nodes. In order to reduce the overhead, some approaches limit the number of copies of message sprayed into network. For example, Spray and Wait [2] assigns a fixed number of messages, hoping at least one copy of them will be forwarded to the destination. However, these algorithms do not pick out the nodes which have the higher probability to meet the destination and forward message to them. Instead, nodes simply send message to every node they meet or a limited number of nodes they encounter.

Mobility pattern has been shown to be useful in link prediction and routing [3-5]. In [3], it is found that the similarity between two individuals' movements strongly correlates with their proximity in the social network. Thus mobility pattern can be used to predict the formation of new links. Moreover, while the short-ranged travel of human is periodic both spatially and temporally and not effected by the social network structure, the long-ranged travel is influenced by social network ties [4]. So the model of human mobility pattern could combine short range movements with travel which is influenced by social network structure. [5] shows the mobility patterns of mobile devices are closely related to users’ social relationships and behaviors, which lays a foundation for applying mobility pattern to routing and forwarding protocols

In order to use the past history of nodes to predict the probability they will meet the destination, some algorithms make use of the human mobility pattern gained from social networks and apply them to the routing for DTN [6,7]. The BUBBLE Rap algorithm [6] uses high centrality nodes and community members of destination as relays. It uses social information including social relationship and mobility pattern to calculate the global ranking and local ranking of nodes in the network then bubbles the message up the hierarchical ranking tree. In [7], the routing protocol for publish-subscribe relationship, in which the network delivers a published message only to nodes who subscribe interests in it, is put forward. The SocialCast algorithm predicts nodes' movement based on the history. Using the prediction, nodes can estimate good message carriers and opportunistically deliver group's main interest to them. Nevertheless, these social-based routing protocols need to frequently update the information gained from social network, such as global ranking and local ranking in BUBBLE Rap and movement prediction in SocialCast, which results in the large overhead.

Compared with related routing protocols for DTN, we believe that our routing protocol which uses the mobility pattern gained from real location data from online social network services, which can efficiently increase the delivery ratio and decrease the network overhead.

3. DESIGN APPROACH

In this section, we present our system design for developing the SGI data collector and SGI routing algorithm. **Fig 2** shows the overview of the implemented modules and how it work with other components. The SGI data collector downloads the social data from Instagram to generate friend relationship data and mobility patterns of users. It then stores these two types of data into our database. The SGI routing algorithm has been implemented in the ONE simulator [8] which already has other existing DTN routing protocols so that we can simply compare the performance between our proposal and other routing protocols on same scenarios. The simulator runs the SGI routing algorithm with the input data from the database. Then it finally gives the results which are the delivery ratio and the network overhead. The following two subsections will describe more details about each module.

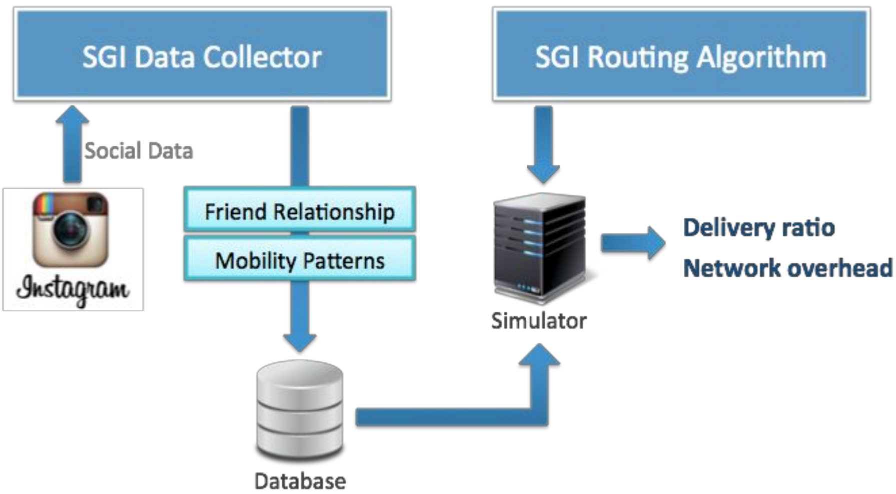

Fig 2. Overview of our system design

3.1. SGI Data Collector

We implemented the SGI data collector by using the Instagram Python API [9] which is provided by Instagram. Our data collector basically requests some information to the Instagram server and receives the data from the server, and this request-response procedure can be divided into three stages. In first stage, the data collector requests recently uploaded photos that have SGI which is UCLA. We decided to get few photos in a specific place as our first stage because a large number of photos from random places will make an user set that are located in multiple places. Since our simulation does not cover large area but we focus on a limited area, for example west side of Los Angeles, we have to filter out if the location of the photo is out of our defined area. After receiving the photos from the Instagram server in the first stage, the data collector retrieve the user ID of each photo. In second stage, the data collector requests user ID of followers of the user set, and the Instagram server gives a list of user IDs to the data collector. We define a friend relationship if a user follows the other user. So that all of the downloaded users are friends of the user who uploaded a photo at UCLA. Finally, in third stage, the data collector requests recent photos of all users in the user set. We filtered out the photos that do not contain SGI or the location of the photo is not inside the area. The remaining photos of each user then represent a mobility pattern of the user. We noticed that high percentage of followers have similar points with its friend which indicate that two friends tend to have a high possibility to meet each other.

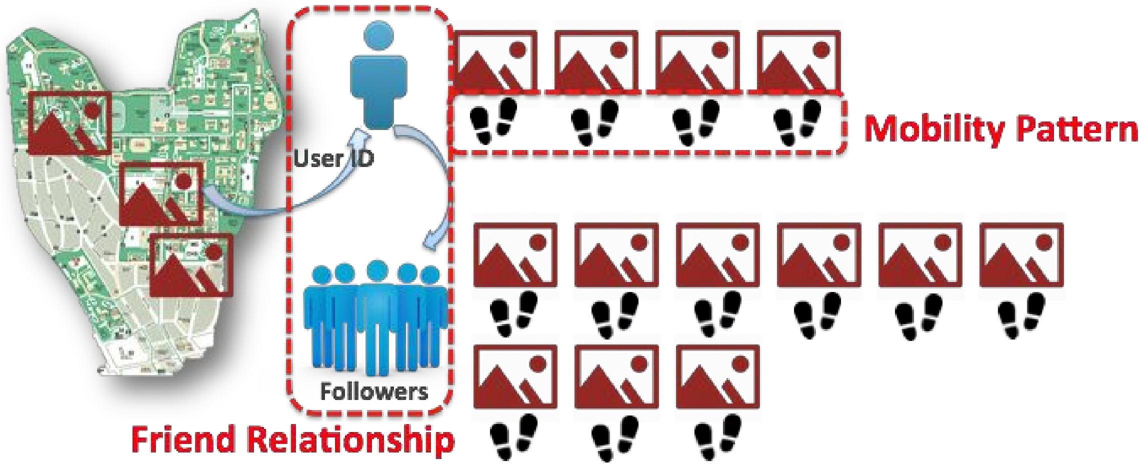

Fig 3. Two types of data that are generated by the SGI data collector

3.2. SGI Routing Algorithm

In this subsection, we explain how the SGI routing algorithm works. There are two assumptions in our proposal. First, every node know its friend nodes and their mobility patterns. We can have this assumption because online social network services provide a list of friends and the places that friends visited. Second, destination node of any packet is a friend of source node. In other words, a node only sends some data to its friends. Based on these two assumptions, the SGI routing algorithm can help node to forward a packet to the destination.

The packet forwarding strategy in SGI routing protocol can be described as 5 steps: (1) a node randomly creates a message to one of its friends. (2) All nodes have their own mobility patterns, so that they move repetitively move upon its mobility pattern. When a node that has the message to forward encounters another node, (3) the node sends the mobility pattern of the destination node to the encountered node. (4) The encountered node then calculates the encounter probability with the destination node. **Fig 4** shows the equation of encounter probability which equals to the number of points that are adjacent to one of the points that destination has divided by the number of combination of mobility pattern of encountered node and that of destination node. This probability represents the chance to meet the destination node by comparing with its own mobility pattern. **Fig 4** also gives an example of encounter probability among three nodes, A, B, and C. In this example, node C has higher encounter probability for node A than node B. That is because node C has 2 common points with node A over total 9 possible points which is about 22%. However, node B has only one common point over 6 possible points which is around 17%. (5) Encountered node sends the calculated encounter probability to the node that gave the destination mobility pattern. Then the node compares the probability with its own probability. If the encounter node has higher probability, then the node copies the message and forwards it to the encountered node in order to have higher delivery probability. Otherwise, the node simply does nothing so that the network overhead can be decreased. These whole procedure from number except (1) is repeated until a node meets the destination node or the message times out.

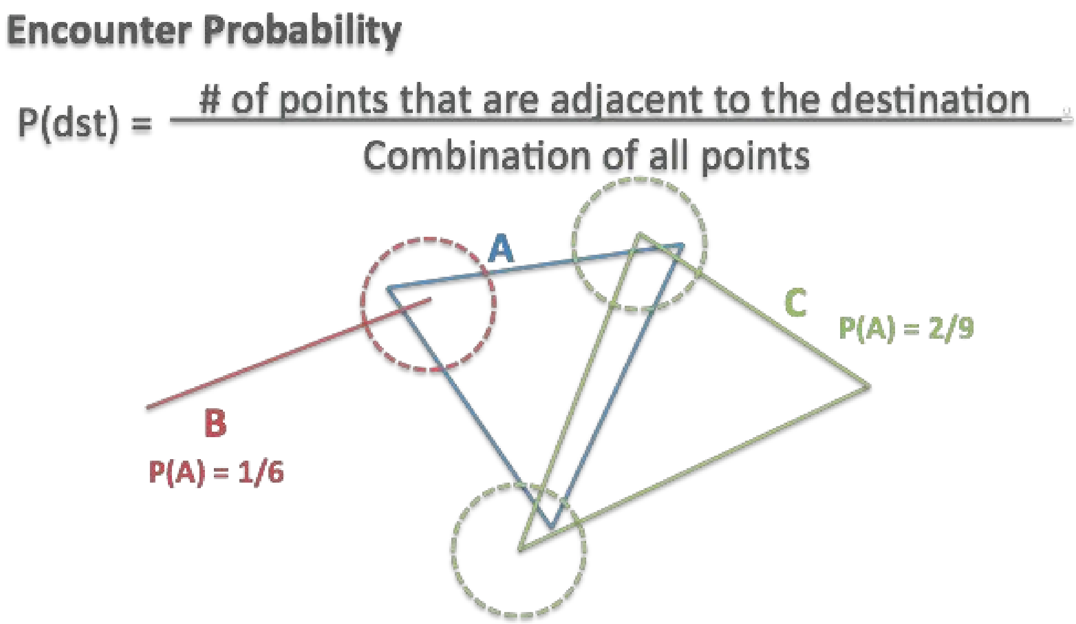

Fig 4. Encounter probability calculation among three nodes

4. PERFORMANCE EVALUATION

4.1. Simulation Environment

In this project, we decided to use the ONE simulator because it already has various features for the DTN simulation. When we run the simulation, the simulator reads the setting text file to set up the simulation environment such as the number of nodes, the number of groups, time-to-live (TTL) value for packets, buffer size, or speed of node groups. Based on this information, the simulator builds the simulation map which is called the world in the ONE simulator, and makes nodes on the world. After running the simulation, the simulator generates one report file that contains many metrics such as delivery probability, delivery latency, network overhead, packet drop ratio, or average number of hops to reach to destinations.

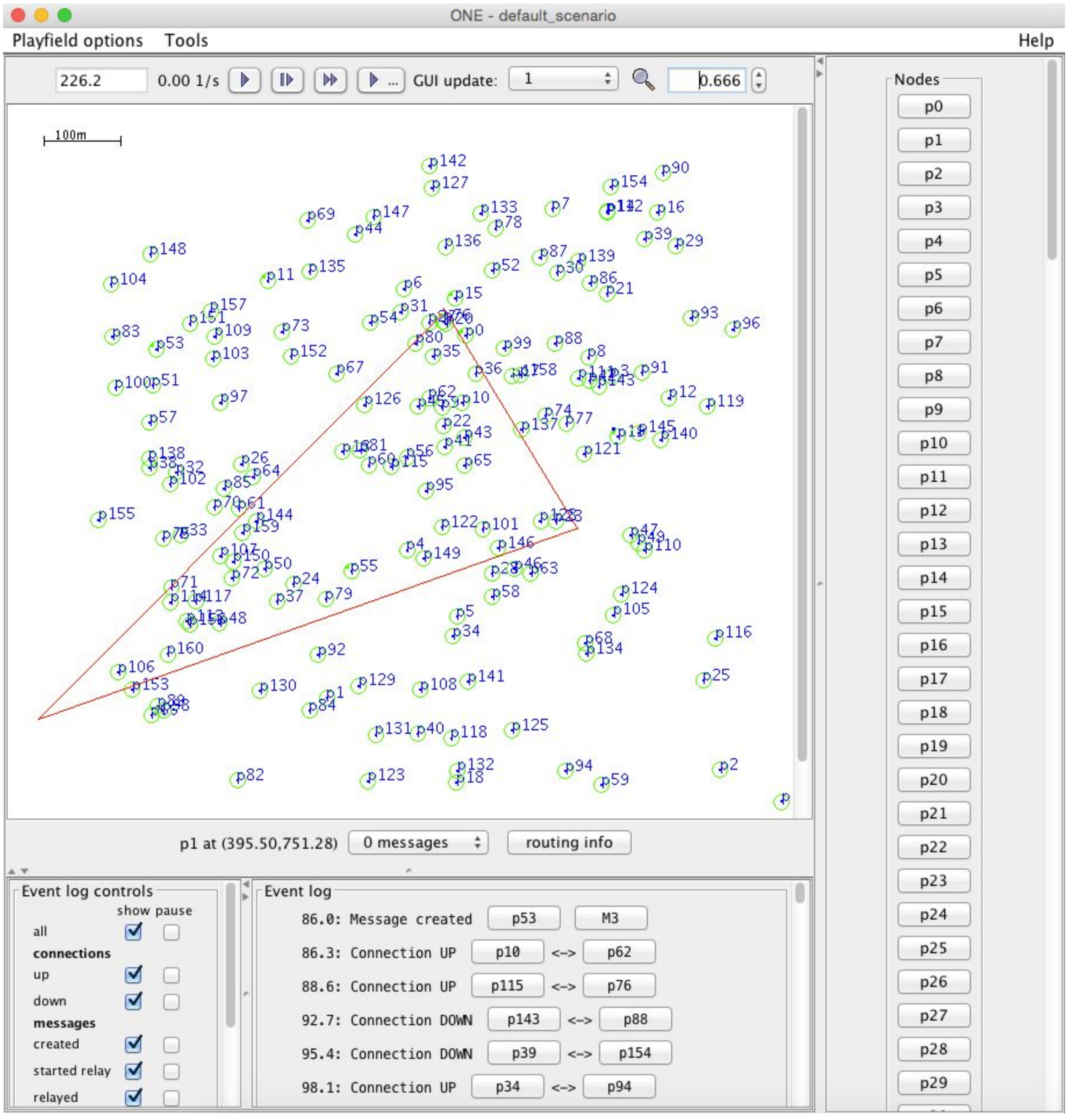

Fig 5. The ONE simulator runs SGI routing protocol

Since we created two types of information which are the friend relationship and the mobility patterns, the simulator has to learn it before running the simulation. We made these two information as two text files and the simulator reads these files whenever it starts simulation. The simulator not only creates the number of nodes but also every node who is aware of its friends and has its own mobility pattern. **Fig 5** shows a screenshot of the ONE simulator that is running the SGI routing protocol. There are 160 nodes in this simulation and each node has a unique ID. The green circle around each node represents the transmission range and the red line in the figure represents the mobility pattern of node p0.

The total number of nodes is 160, and 80 of them are come from the Instagram data so that they have friend relationships and generate packets to its friends and also have pre-defined mobility patterns. The other 80 nodes are random node which randomly choose next waypoint and do not generate any packets. These random nodes only help forward packets to destination as intermediate nodes. The buffer size of each node is 50 Mbytes, and each packet size is 0.5~1.5 Mbytes. We have changed packet TTL values to evaluate performance. Each simulation time is 10 hours, and we have run 10 times for each condition then get the average value for the result.

4.2. Simulation Results

We compared three different DTN routing protocols in our simulation; SGI, Spray-and-Wait, and Epidemic. In order to evaluate the performance, we checked each routing protocol's delivery ratio and the network overhead. The delivery ratio equals to the total number of delivered packets divided by the total number of packets. Since it only cares the number of delivered packets, the number of copies does not matter in this case. However, the network overhead equals to the number of relayed packets minus the number of delivered packets, divided by the number of delivered packets. It indicates the ratio of number of redundant packets in the network.

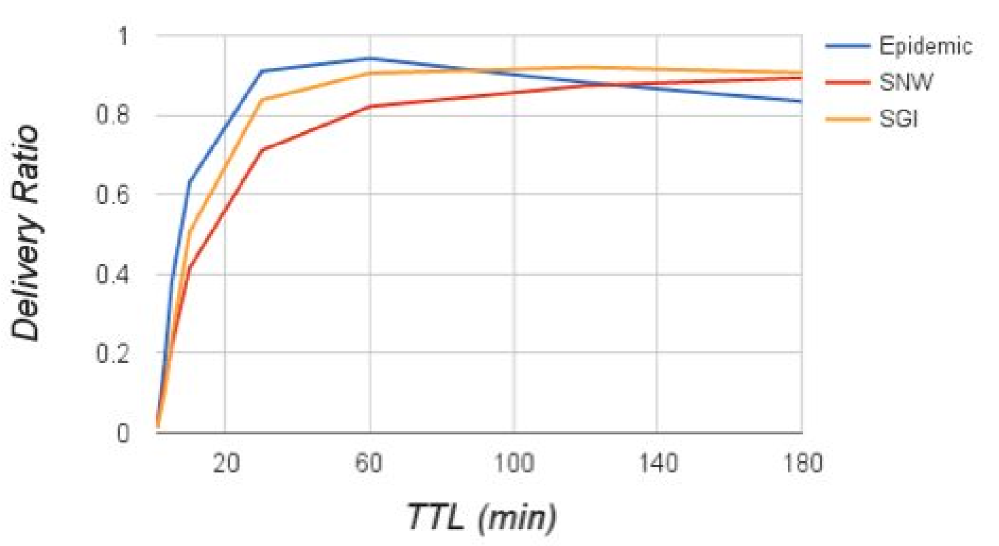

Fig 6. Delivery ratio over TTL changes

Fig 6 shows the delivery ratio of three routing protocols. When the ttl value is less than 80 minutes, Epidemic routing protocol has the highest delivery ratio among three protocols. However it decreases while the ttl value is increasing. That is because Epidemic generates too many copies in the network, so if the ttl is large then the packets are more likely not dropped by time out but they are dropped because of overflowing the buffer in each node. However, our proposed protocol limits the number of copies by calculating the encounter probability so that it has higher delivery ratio when the TTL value is large. The Spray-and-Wait protocol always has lower performance in our simulation because it only propagates five copies into the network by using the binary mode.

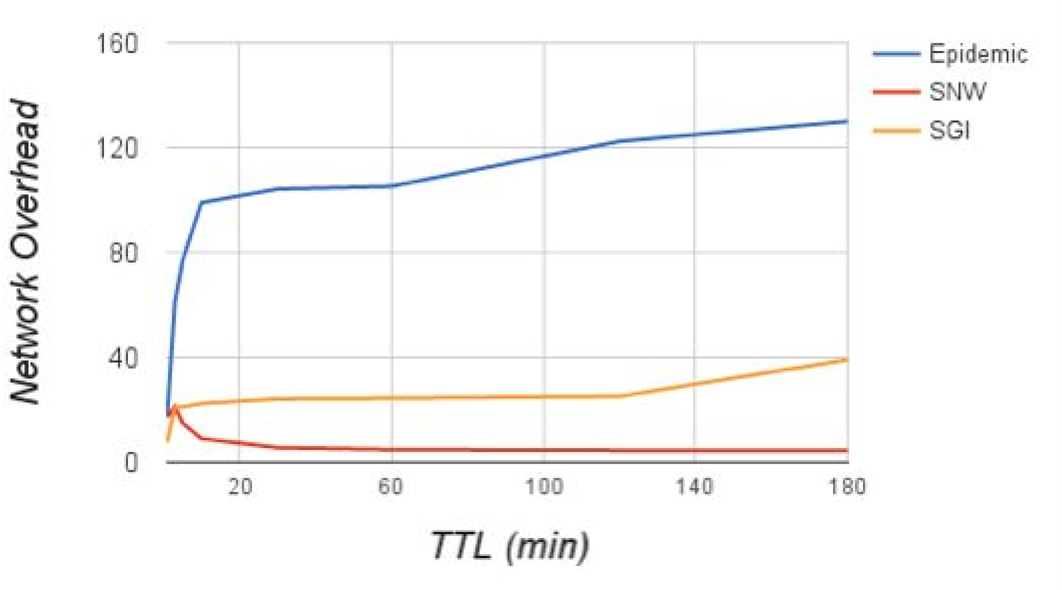

Fig 7. Network overhead over TTL changes

Fig 7 shows the network overhead. The network overhead tends to increase when the TTL value increases except the Spray-and-Wait protocol. That is because the Spray-and-Wait protocol always limits the number of copies. The SGI protocol significantly decreases (about 80%) the network overhead compared with Epidemic routing protocol.

By running the simulation, we could check that our proposal has relatively higher delivery ratio than existing routing protocols and also decreases the network overhead significantly.

5. CONCLUSION AND FUTURE WORK

In this paper we presented a social geo-location information based routing protocol for DTN, which uses mobility pattern gained from online social network services to forward message to nodes who are more likely to meet with destination. We firstly builded the SGI data collector to download social information from Instagram and generated the friend relationship and the mobility patterns based on it. Then we simulated our proposed SGI routing protocol, in which nodes calculate the probability that encountered node will meet with destination and forward message to those with higher probability, by using the ONE simulator. Simulation results show that the proposed SGI routing protocol has relatively higher delivery ratio and lower network overhead compared with Spray-and-Wait and Epidemic routing protocol.

We believe that this paper represents a first step in using real location data from online social network services to build mobility pattern and proposing a routing protocol for this situation. In the future, we plan to test and modify the SGI routing protocol in more complicated scenarios where nodes have more random mobility pattern in a larger area. This will facilitate the application of the SGI routing protocol in real world scenario.

REFERENCES

[1] A. Vahdat, D. Becker, et al. Epidemic routing for partially connected ad hoc networks. Technical report, 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 ACM SIGCOMM workshop on Delay-tolerant networking, pages 252–259, 2005.
[3] Wang D, Pedreschi D, Song C, et al. Human mobility, social ties, and link prediction[C]//Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011: 1100-1108.
[4] Cho E, Myers S A, Leskovec J. Friendship and mobility: user movement in location-based social networks[C]//Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011: 1082-1090.
[5] Xia F, Liu L, Li J, et al. Socially aware networking: A survey[J]. 2013.
[6] Hui P, Crowcroft J, Yoneki E. Bubble rap: Social-based forwarding in delay-tolerant networks[J]. Mobile Computing, IEEE Transactions on, 2011, 10(11): 1576-1589.
[7] Costa P, Mascolo C, Musolesi M, et al. Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks[J]. Selected Areas in Communications, IEEE Journal on, 2008, 26(5): 748-760.
[8] Keränen, A., Ott, J., & Kärkkäinen, T. (2009, March). The ONE simulator for DTN protocol evaluation. In Proceedings of the 2nd international conference on simulation tools and techniques (p. 55). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
[9] Instagram Python Library, “https://github.com/Instagram/python-instagram”