6-2010

A social transitivity-based data dissemination scheme for opportunistic networks

Jaesung KU

Yangwoo KO

Jisun AN
Singapore Management University, jisunan@smu.edu.sg

Dongman LEE

Citation

KU, Jaesung; KO, Yangwoo; AN, Jisun; and LEE, Dongman. A social transitivity-based data dissemination scheme for opportunistic networks. (2010). *Proceedings of 2010 IEEE International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), Montreal, Canada, June 14-17*. 1-9.
Available at: https://ink.library.smu.edu.sg/sis_research/6765

A Social Transitivity-based Data Dissemination Scheme for Opportunistic Networks

Jaesung Ku, Yangwoo Ko, Jisun An, and Dongman Lee

Department of Computer Science, KAIST
119 Munji-ro, Yuseong-gu, Daejeon, Republic of Korea
{kujaesung, yko, jisun, dlee}@kaist.ac.kr

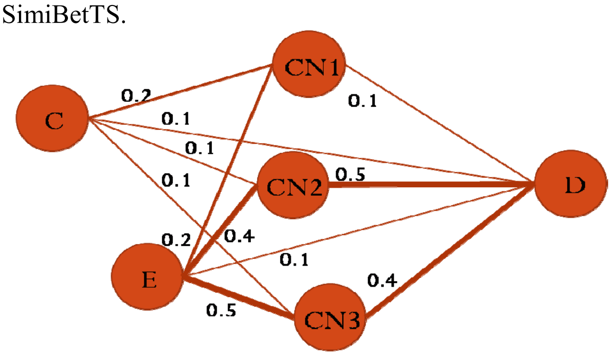
Abstract— **A social-based routing protocol for opportunistic networks considers the direct delivery as forwarding metrics. By ignoring the indirect delivery through intermediate nodes, it misses chances to find paths that are better in terms of delivery ratio and time. To overcome this limitation, we propose to incorporate transitivity, which considers the indirect delivery through intermediate nodes, as one of the forwarding metrics. We also found that some message forwards do not improve the delivery performance. To reduce the number of these useless forwards, the proposed scheme forwards messages to an encountered node when the increase of total utility value is greater than a threshold. Using a simulator with real world trace data sets, we compare the proposed scheme with the existing protocols, epidemic routing and SimBetTS. Compared with SimBetTS, the proposed scheme increases delivery ratio by 1.5 percent and decreases delay time by 2 percent while reducing overhead by 30 percent.**

Keywords— **Opportunistic Routing, Social-based Routing, Data Dissemination, Delay-Tolerant Networks, Opportunistic Networks**

I. Introduction

Opportunistic networks [1], also called intermittently connected networks or delay-tolerant networks (DTNs), aim at delivering messages even in a disconnected environment. As they move around, nodes are exchanging messages with their neighbors which have a better chance of bringing the messages to destination nodes. Epidemic message exchange is proposed in the first place. It delivers a message to a destination but gives a rise to a large amount of message copies in the network which may lead to network congestion. In order to deliver messages more efficiently, several schemes [17-23] have been proposed.

One of such efforts is to consider social perspectives of users' mobility: social-based opportunistic network routing scheme [3, 4, 5]. HiBOp[3] exploits high level context like user's residence and work location for learning how users relate with each other and optimizing data forwarding. Bubble Rap[5] adopts two social metric, community and centrality for selecting efficient forwarding paths. SimBetTS[4] incorporates as its routing metrics tie strengths as well as community and centrality. It exploits the total utility value, which is derived from three social metrics, similarity, betweenness centrality, and tie strength, to determine whether to forward messages to an encountered node. Though the existing schemes including SimBetTS assume the indirect delivery in which a node delivers messages to a destination node through intermediate nodes, their forwarding metrics only consider the direct delivery. For example, a node in the same community with a destination node is likely to meet the destination node than others in different communities. Similarity is considered as a metric to detect community. And tie strength in SimBetTS expresses how well a node can deliver messages to a destination node by itself. Although betweenness centrality allows messages to be forwarded to popular nodes that have a high chance to meet other suitable carriers, it is used for preventing message from being isolated in SimiBetTS.



(a) Ties with strength in real networks

Node	Similarity	Betweenness Centrality	Tie Strength	Total Utility Value
C	3	2.5	0.1	0.5
E	3	2.5	0.1	0.5

(b) Calculation of total utility value

C: current node, E: encountered node, D: destination node
CN[1-3]: common neighbor nodes between C and D, E and D
Thickness and number of lines: tie strength

Fig. 1 Ties with strength in real networks and calculation of total utility value

Let us assume that tie strengths between nodes are as shown in Figure 1(a). Based on this, Figure 1(b) shows how total utility values are calculated when node C encounters node E. Both node C and node E have the same number of neighbor nodes to node D so that their similarity, betweenness centrality and tie strength are the same. Thus, since SimBetTS considers both node C and node E have the same delivery probability, node C does not forward messages to node E. As shown in Figure 1(a), social distance [16] from E to D can be longer than the sum of those from E to CN2 and CN2 to D.

That is, the direct delivery probability from E to D can be less than the indirect delivery probability from E to D through an intermediate node, CN2. When considering the indirect delivery probability, we can easily find that node E has higher chances to deliver the messages to node D than node C.

In this paper, we propose a new scheme that exploits not only two metrics, betweenness centrality and tie strength, proposed by SimBetTS but also an additional social metric, social transitivity. The proposed scheme defines social transitivity as the degree of delivery probability indirectly through intermediate nodes. Transitivity is calculated from the sum of product of tie strengths between nodes in the all paths from a sender node to a destination node. The proposed scheme combines transitivity and tie strength for calculating total utility value.

When a node meets another node having a greater total utility value for a message, the message is forwarded. If the difference of total utility values of the two nodes is very small, the forwarding may incur overhead without improving the delivery performance. We propose to set a threshold for the difference of total utility values. If the difference is below the threshold, even if an encountered node has a greater total utility value than the current node, the current node does not forward messages to the encountered node unconditionally. In the proposed scheme, a node derives the threshold from the average of differences of total utility values per hop for all delivered messages to it.

We implement an event-driven simulator and compare the proposed scheme with the existing works using real world trace data sets, Infocom05 [9] and Cambridge [15]. The proposed scheme outperforms SimBetTS with the increased delivery ratio by 1.5 percent and the decreased delay time by 2 percent while reducing overhead by 30 percent. The proposed scheme achieves delivery ratio and delay time close to those of epidemic routing [2], while it incurs far less overhead.

The rest of the paper is organized as follows: In Section II, we provide comparison of existing schemes and especially, gives more detailed information about SimBetTS in terms of social metrics, such as similarity, betweenness centrality, and tie strength. Section III explains the design considerations for the proposed scheme. Then, in Section IV, we propose a new social metric, transitivity, and show how to calculate the total utility value based on transitivity, and introduce the routing algorithm. In Section V, we present simulation results and an evaluation for our proposed scheme. Finally, Section VI gives conclusion and describes future work.

II. Related Work

HiBOp [3, 6] uses two social metrics: community and tie strength. Community is calculated based on the match, a measure of similarity, between the context of a destination node and a sending node. It shows that a node, which lived or commuted in near areas with a destination node, has more chances to deliver the messages. Tie strength identifies links which have a higher probability of availability. It is possible to quantify tie strength between two nodes based on how frequently they encounter each other, how recently they have met, and how long they were together.

SimBet [7] considers both centrality and community as metrics to determine the next forwarding nodes. In addition to these two metrics, SimBetTS [4] incorporates tie strength to address the problem of SimBet's unawareness of the time-varying nature of link availability.

Like SimBet, Bubble Rap [5, 8] also uses centrality and community to calculate the delivery probability. Each node forwards messages to a more popular node until the messages are delivered to a node which belongs to the same community with the destination node. Then, the messages are forwarded to a more popular node within the community until delivered to the destination node.

Social metrics used in the schemes above are derived from the properties of social networks, which are examples of complex networks [10]. There are various properties of social networks, such as the small-world phenomenon [12], transitivity, and degree distribution [11].

TABLE I
Social Network Properties and Metrics in Social-based Schemes

Schemes		HiBOp	Bubble Rap	SimBet TS
Social metrics	Community	v	v	v
	Tie Strength	v	-	v
	Centrality	-	v	v
Social networks properties	Community structure	v	v	v
	Degree distribution	v	v	v
	Short average path length	v	v	v
	Degree correleation	v	-	v

In table I, we show social metrics and social network properties which are used in the above social-based schemes. There are four social network properties: community structure, degree distribution, short average path length, and degree correlation. Degree distribution describes how many links a node has to other nodes. Community structure means that two nodes within same community have a higher density of links than different communities. Short average path length is similar to the small-world phenomenon. Degree correlation shows that some selective links, which is called assortative mixing or homophily, exist in social networks. Based on the social network properties, three social metrics, community, centrality, and tie strength, are derived and adopted to calculate the total utility value which represents the links probability and is used to determine whether to forward messages to encountered node.

III. Design Considerations

In this section, we extract the key requirements necessary to incorporate social transitivity into calculating of total utility value and to determine whether a forward is useful or not.

The word, transitivity, refers to the situations where mutual links exist between three nodes in a network [14]. Two separated nodes can make a connection through a mutually shared node. If one mutually shared node can increase the likelihood of a connection between two nodes, multiple mutual nodes should boost that probability even more. Transitivity is dependent on the quality of links as well. In other words, transitivity of a pair of nodes depends on how many mutually shared nodes a couple of nodes have as well as on how strongly they are connected with shared nodes [13]. Thus, the proposed scheme should consider the tie strengths of multiple mutually shared nodes when calculating total utility value. If all intermediate ties within a path to the destination node are strong, the delivery probability can be increased. When the number of paths increases, it also increases delivery probability. Because the paths are independent of each other, transitivity should derive the effect from the number of paths through sum of it.

In the existing schemes, whenever encountering nodes which have a greater total utility value of messages, a node forwards the messages to the encountered nodes. However not all forwarding results in the delivery of the messages to the destination nodes. For reducing the number of useless forwards, the proposed scheme should provide the criterion for determining whether the forward is useless or not. A node should know how much difference of the total utility values can assure that the forward is not useless From the hop counts and difference of total utility value of the delivered messages to a destination node, we can set the minimum bound of the difference of total utility values that assures the delivery.

IV. PROPOED SCHEME

A. Overview

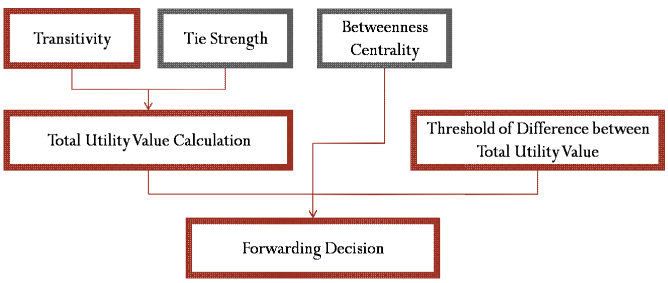


Fig. 2 Overview of the proposed scheme

The proposed scheme incorporates transitivity into the calculation of total utility value in addition to tie strength used in the exiting work. In forwarding decision, threshold of difference between total utility values is used for reducing the useless forwards. Unlike the existing work, betweenness centrality is not incorporated in the calculation of total utility value and is used only when a destination node is not known to both a current node and neighbors of it to avoid messages being isolated.

B. Calculation of Transitivity

The proposed scheme uses transitivity to consider the indirect delivery probability through a group of intermediate nodes. Although a node does not have a fully acceptable link quality to a destination node, the neighbors of the current node can support strengthened indirect delivery probability to the destination node.

A node can calculate its transitivity to a destination node from the strength of ties (1) between the nodes and its neighbors, (2) between the node's neighbors, and (3) between the node's neighbors and the destination node. Each node maintains the tie strength table which stores the strength of ties of between itself to all encountered nodes as well as those between other nodes. When a node encounters another node, both nodes exchange their tie strength tables. Through this process, a node obtains tie strength information for other nodes some of which the node has not encountered.

For example, as shown in figure 3, node C has *n* neighbor nodes, N1 through Nn. By exchanging tie strength table, node C knows the tie strengths not only between itself and node N1 through Nn, but also the tie strengths between node N1 through Nn and a destination node D. Node C's transitivity to node D, denoted as $T_C(D)$, is calculated using formula (1).

$$T_C(D) = \sum_{j=1}^{n} S_C(Nj) \times S_{Nj}(D) \qquad (1)$$

where $S_x(y)$ represents the tie strength between node x and node y.

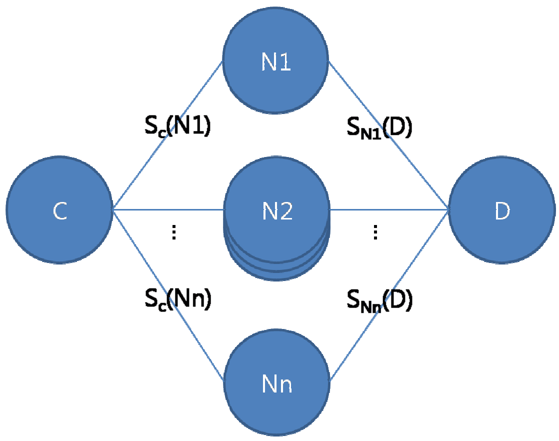


Fig. 2 Transitivity calculation

Tie strength is derived from combining the three components: frequency, duration, and recency. The calculation methods are the same as in SimBetTS [4].

C. Calculation of Total Utility Value

When node A encounters node B, both nodes decide which message should be forwarded to the other node. A message that node A has in its buffer while node B has not is forwarded to node B if the total utility of node B regarding the destination of the message is greater than that of node A. In the proposed scheme, the total utility value is obtained by relative sub-utility value of tie strength and transitivity. Unlike SimBetTS, it does not incorporate similarity and betweenness centrality in the utility calculation and the rationale behind this is given in section V.*B*.

When node A meets another node B, its total utility value for a message destined to node D, $U_A(D)$, is derived from the weighted average of the relative transitivity utility denoted as $RT_A(D)$ values and the relative tie strength utility denoted as $RS_A(D)$. We use the weighed factor as α, and both two attributes, transitivity and tie strength, are considered to be of equal importance in the simulation. In short, the total utility value of node A for the message destined to node D is calculated as

$$U_A(D) = \alpha \times RT_A(D) + (1 - \alpha) \times RS_A(D) \qquad (2)$$

where

$$RT_A(D) = \frac{T_A(D)}{T_A(D) + T_B(D)} \qquad (3)$$

$$RS_A(D) = \frac{S_A(D)}{S_A(D) + S_B(D)} \qquad (4)$$

Please note that all utility values are relative metrics that are defined in comparison with an encountered node B. As shown in the formula (2, 3), we cannot calculate the relative utility values if both node A and B has zero transitivity utilities and tie strengths to a destination node. This is a case where node encounters are sparse and the destination node has a small number of social interactions. The proposed scheme uses betweenness centrality, which is independent of a destination node, to handle this situation. In such a case, the message is forwarded to node B if the betweenness of node B is greater than that of node A.

D. Calculation of Threshold

It is reasonable to conjecture that the delivery probability will not increase that much when a node forwards messages to an encountered node that has slightly higher total utility value than it. We need to know how big the increment of total utility value should be to assure that a message forward is not wasted. We exploit the history of successful message deliveries as a base for such a decision. A node forwards a message to an encountered node when the total utility value of the encountered node is bigger than that of itself. The differences between the total utility values vary across hops as well as destinations. Equipped with the statistics of such differences, a node can tell whether the total utility value of an encountered node regarding a destination is relatively big enough. There are an issue in the exploitation of overall statistics of message deliveries in forwarding decisions. Each node cannot access delivery statistics of all other nodes. The proposed scheme tackles the issue by exchanging delivery statistics with encountered nodes.

Based on the delivery information obtained locally as well as from other nodes, each node calculates the threshold value independently by taking the arithmetic average of differences of total utility values of all forwards involved in all successful deliveries. After calculating the threshold, a node forwards messages to the encountered node if the difference between the total utility values for two nodes is greater than the threshold. In case of the difference is less than the threshold, the node determines whether to forward or not based on the ratio of the difference between total utility values to the threshold, which is shown in (5). The node forwards messages to the encountered node in the probability of the ratio.

$$\text{Ratio} = \frac{\text{Difference of total utility value}}{\text{Threshold}} \qquad (5)$$

E. Forwarding Decision

Table II shows the routing algorithm when node A encounters node B. When node A receives a HELLO message from node B, node A updates the tie strength for node B and any messages destined for node B are sent. The two nodes then exchange both encounter vectors and tie strength vectors that are used for updating the betweenness centralities and tie strength vectors. Both nodes also update the threshold by exchanging its delivery statistics. Subsequently, both nodes exchange a summary vector of messages, which includes destination nodes, and betweenness centrality that they are currently carrying.

TABLE II
ROUTING ALGORITHM, PSEUDOCODE FOR NODE A

```
PROCEDURE Handle-HELLO(node B)
  FOREACH m in Messages DO
    IF m.destination == B
      THEN DeliverMessage(m, B)
  ENDFOR
  ExchangeEncounterVectors(B)
  ExchangeTieStrengthVectors(B)
  UpdateBetweennessCentrality()
  UpdateTieStrength()
  sv = ExchangeMessageSummaryVector(B)
  ExchangeMessages(B, sv)
END

PROCEDURE ExchangeMessages(node B,
SummaryVector sv)
  MessagesToRequest = {}
  FOREACH destination D in sv DO
    Calculate TA(D), TB(D)
    IF({TA(D), TB(D), SA(D), SB(D)}=={0})
    THEN
      IF(BetweennessCentrality(A)>BetweenessCentrality
      (B) )
      THEN MessagesToRequest.Add(D)
    ELSE
      Difference = UA(D) - UB(D)
      IF (Difference > Threshold)
        THEN MessagesToRequest.Add(D)
```

```
        ELSE
          IF((Difference>zero)&&(Math.random()<Ratio))
            THEN MessagesToRequest.Add(D)
   ENDFOR
END
```

V. Simulation

In this section, we describe our simulation setup and results. We then explain the performance of the proposed scheme by comparing it with SimBetTS and Epidemic routing schemes. We implement a custom event-driven simulator. In case of Epidemic routing, which uses random pair-wise exchanges of messages among nodes, we suppose that both encountered nodes can receive all unseen replicas of messages like in a flooding-based scheme.

A. Simulation Environments

The trace files used for the event-based simulation are from Infocom05 [9] and Cambridge [15].

We generate 5 sets of scenarios each of which contains a set of message delivery requests. Each delivery request is tuple of (sender node, destination node, sent time). Nodes are selected from all nodes in the traces files using a uniform random distribution. All simulation results are the average of these 5scenarios.

1) *Simulation with Infocom05 trace*: The number of whole participants and devices are 41 and 264, respectively, and the experiment lasted three days. A contact is considered as a symmetric interaction so that both nodes can exchange its own messages with each others. The contact duration is the difference between beginning time and ending time of the contact. Table III shows the summary of simulation environments. Buffer size is unlimited and Time-to-live (TTL) is set as 86,400 seconds. And the whole simulation time is 274,883 seconds. The first message sent time is 27,184 seconds after the beginning of simulation. Then, until 239,584 seconds, all 41 nodes generate 20 messages which towards different destination nodes per hour, totally 49,200 messages. The weight factor for tie strength and transitivity is 0.5. And two versions of SimBetTS and the proposed scheme are simulated respectively: a single-copy version and a multi-copy version.

2) *Simulation with Cambridge trace:* The experiment, Cambridge, lasted five days and the number of all participants and devices are 12 and 223, respectively. The whole simulation time is 454040 seconds and the first message sent time is the time after 2613 seconds from the beginning time of the simulation. Then, until 400000 seconds, all 12 nodes generate 10 messages which towards different destination nodes per hour, for a total of 13200 messages. Unlike Infocom05, because of small sizes of contacts and number of nodes, a single-copy version of SimBetTS and the proposed scheme is simulated.

TABLE III
Summary of Simulation Environments

Parameter	Infocom05	Cambridge
Simulation time	274883(s)	454040(s)
Number of sent messages	49200	13200
Number of nodes	41	12
Time-to-live	86400(s)	
Buffer size	Infinite	
Schemes	SimBetTS(R1, R4), Epidemic routing, Proposed(R1, R4)	
Weight factor(α)	0.5	

B. Evaluation

In this section, we show four metrics for comparing the delivery performance of the schemes. We then compare the proposed scheme with SimBetTS and Epidemic routing based on these metrics.

1) *Evaluation metrics*: We use four metrics for evaluation: delivery ratio, delay time, overhead, and hop count. Delivery ratio means the percentage of the number of delivered messages to destination nodes among total number of sent messages from source nodes. Delay time is the average elapsed time for delivered messages to destination nodes. Overhead presents the amount of generated message traffic to deliver a message. In other words, this is computed by dividing the total number of forwarded messages by the number of delivered messages. Finally, we can get the information of message path length from the hop count which is the average hop counts for total delivered messages.

2) *Performance Comparison using Infocom05 trace:* Figure 4 shows the results for all delivered messages. As expected, Epidemic routing achieves the greatest message delivery performance, about 43000 delivered messages. The proposed scheme performs better than SimBetTS. While similarity just takes a message to the same community with a destination node, and tie strength provides the direct delivery probability, transitivity complements them by incorporating an indirect delivery probability to the message. When analyzing the trace data set, some time periods, such as 120000s to 150000s or 200000s to 240000s, have only a small number of contacts compared to other time periods. So, the number of delivered messages is not increased during these periods. Conversely, because of the large number of contacts, the number of delivered messages climbs at a fast pace in other time periods.

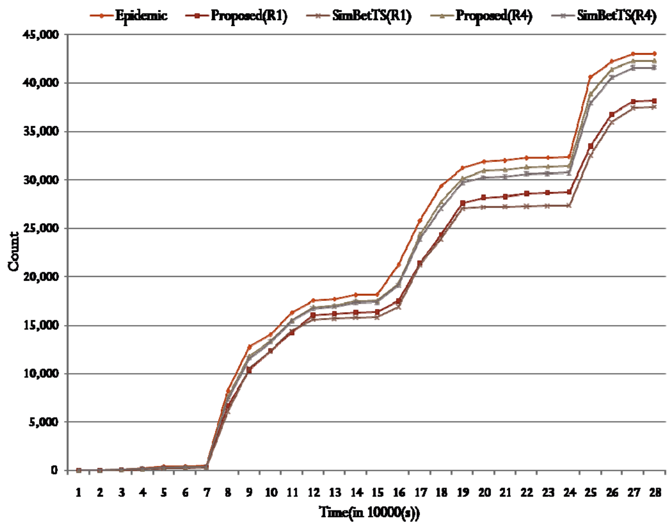


Fig. 3 The total number of delivered messages for Infocom05 trace

In figure 5, the whole delivered messages are divided by the delay time every 3000 seconds. Epidemic routing provides smaller delay time than other schemes. The ultimate goal of our scheme is to have similar performance to that of Epidemic routing. Compared to SimBetTS, the proposed scheme is becoming more like that of Epidemic routing by being shift left. In SimBetTS, as the simulation goes on, there is nothing to choose among the similarities of all nodes. Thus, determining whether to forward or not depends on the tie strength and betweenness centrality. These two metrics make messages on a node delayed, because the node just waits until encountering either the destination node or a node which has a higher total utility value than itself. When analyzing the path of all delivered messages, we can easily see the examples for the above case that cause message delay. However, transitivity, which exhibits the indirect delivery probability through neighbor nodes, could avoid the situation where messages are delayed.

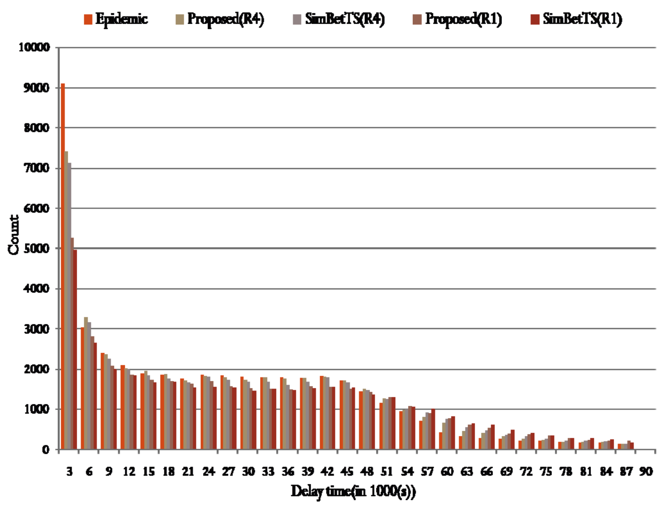


Fig. 4 Frequency distribution for the number of delivered messages for Infocom05 trace

Table IV summarizes the simulation results. As mentioned before, the proposed scheme shows better performance in delivery ratio and delay time than SimBetTS. Although the proposed scheme uses transitivity, which involves the indirect delivery probability through neighbor nodes and causes the number of forwards to increase, the threshold reduces the amount of forwards. Figure 6 shows the total number of forwards, regardless of whether the message is delivered to the destination node or not.

Use of multi-copy replicas in the proposed scheme effectively reduces performance gap between the proposed scheme and the epidemic routing except for the deficiency of messages having a short delay time, shown like the first column in figure 5. This problem remains as future work.

TABLE IV
COMPARISON OF SIMULATION RESULTS FOR INFOCOM05 TRACE

Schemes	**Delivery Ratio (%)**	**Delay Time (s)**	**Overhead**	**Hop Count**
Epidemic	87.6	23132	36.1	3.50
Proposed(R4)	86.0	25212	9.41	2.64
Proposed(R4)-global	86.1	25178	9.42	2.61
SimBetTS(R4)	84.5	25596	12.4	2.94
Proposed(R1)	75.5	27670	3.12	2.71
Proposed(R1)-global	75.7	27680	3.20	2.72
SimBetTS(R1)	74.1	28362	4.31	3.69

Unlike SimBetTS, similarity is replaced by tie strength and betweenness centrality is used conditionally.

As mentioned in section I, similarity exhibits the probability of being in the same community among nodes so that within the same community, there are more frequent interactions among neighbor nodes than between different community. When considering contact information between two specific nodes, tie strength could provide a fine-grained direct delivery probability, which overcomes the limitation of similarity.

Betweenness centrality performs well when the destination node of a message is not known to a sending node or an encountered node. Under this situation, the message could be forwarded to a popular node which has a higher chance to meet other suitable carriers so that it keeps the message from being isolated. However, in SimBetTS, which combines three social metrics, betweenness centrality is misused except for preventing the message from being isolated. For example, a current node has high tie strength for a destination node with low betweenness centrality. When an encountered node has much higher betweenness centrality than the current node, a message to the destination node has no choice but to be forwarded to the encountered node. This example shows that betweenness centrality is likely to causes the message to be delayed. Thus, the proposed scheme uses betweenness centrality conditionally only in the case of a message being isolated.

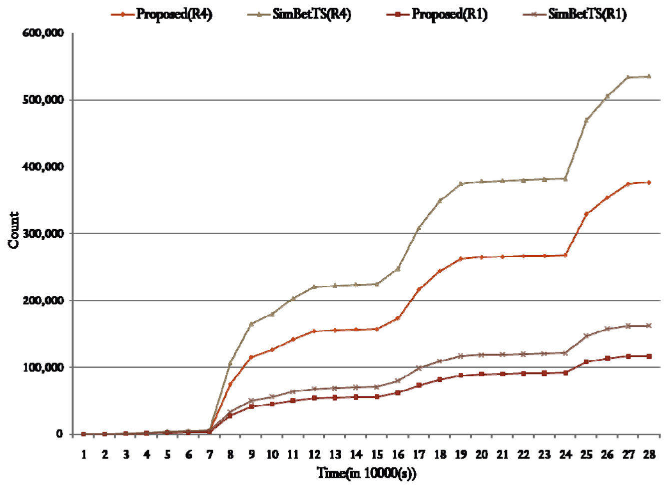


Fig. 5 The total number of forwards for Infocom05 trace

Compared with the delivery performance of global knowledge based proposed scheme, that of the estimated knowledge gives comparable results. Because, In figure 7, as time goes on, the threshold from the distributed method which uses statistical approximates to that of global knowledge.

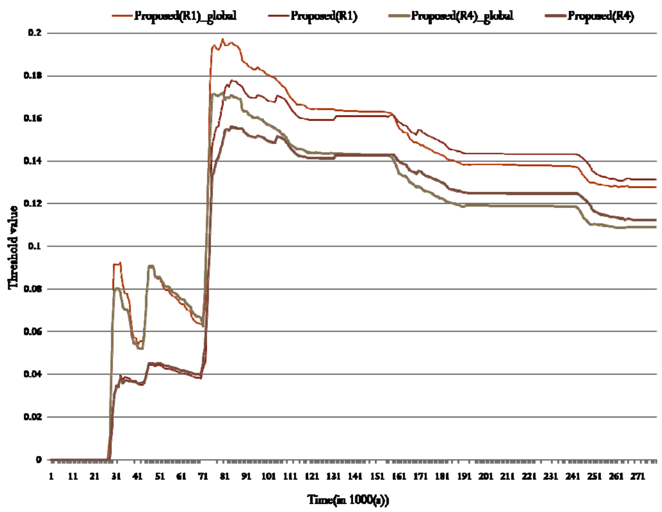


Fig. 6 The threshold for Infocom05 trace

3) *Performance Comparisons using Cambridge trace:* Due to the small sizes of contacts and the number of nodes, the delivery performance of SimBetTS and the proposed scheme could not be evaluated appropriately, especially, in case of a multi-copy version. So, table 5 shows the summary of simulation results for a single-copy version of SimBetTS and the proposed scheme. The propose scheme gives slightly better delivery performance than SimBetTS.

TABLE V
COMPARISON OF SIMULATION RESULTS FOR CAMBRIDGE TRACE

Schemes	Delivery Ratio (%)	Delay Time (s)	Overhead	Hop Count
Epidemic	54.6	29591	19.28	2.52
Proposed(R1)	48.3	30741	2.93	2.32
Proposed(R1)-global	48.4	30657	2.89	2.40
SimBetTS(R1)	46.7	30890	3.13	2.37

As shown in figure 8 and figure 9, Epidemic routing achieves the best performance for delivery ratio and delay time, but remains the largest overhead. Performance metrics of the proposed scheme are located between those of Epidemic routing and those of SimBetTS. As considering the indirect delivery probability through neighbor nodes, nodes enlarge the view to deliver messages to destination nodes so that the proposed scheme could increase the delivery probability with decreased delay time.

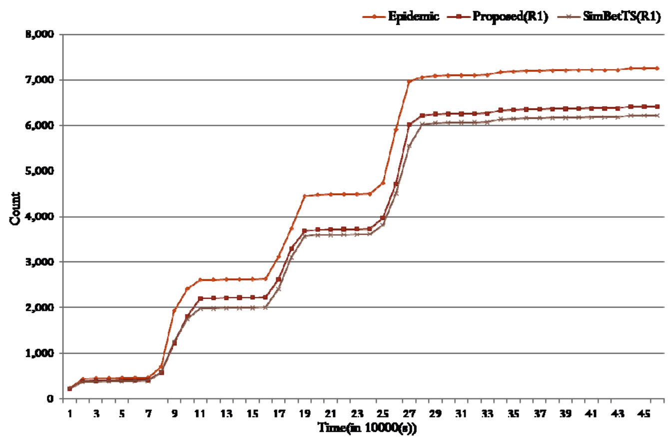


Fig. 7 The total number of delivered messages for Cambridge trace

In figure 8, there are highs and lows in delivery performance on the sections, because the number of contacts between nodes is different for each section.

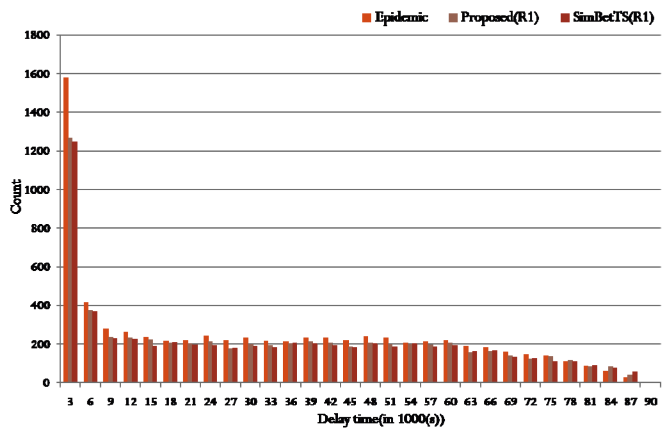


Fig. 8 Frequency distribution for the number of delivered messages for Cambridge trace

Figure 10 exhibits the total number of forwarded messages. The number of forwards for Epidemic routing is too big to fit in a graph and it is omitted. Even if the proposed scheme consider the indirect delivery probability that increases overhead, due to the use the threshold, the number of forwards in the proposed scheme is less than that of SimBetTS.

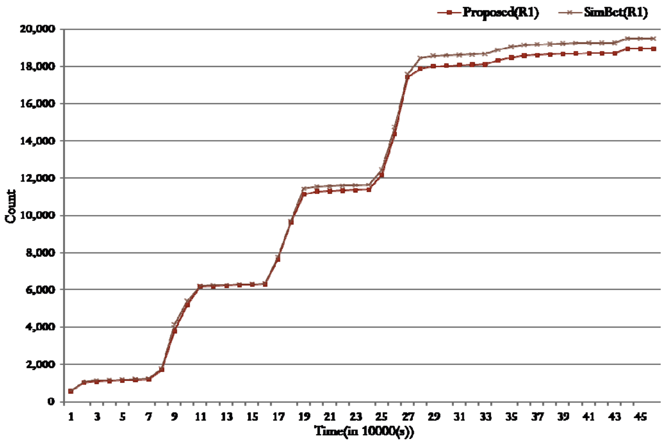


Fig. 9 The total number of forwards for Cambridge trace

Figure 11 shows that as time goes on, the threshold from the distributed method using statistical approximates is following that from global knowledge. Thus, the performance from estimated knowledge provides comparable results to that of global knowledge based.

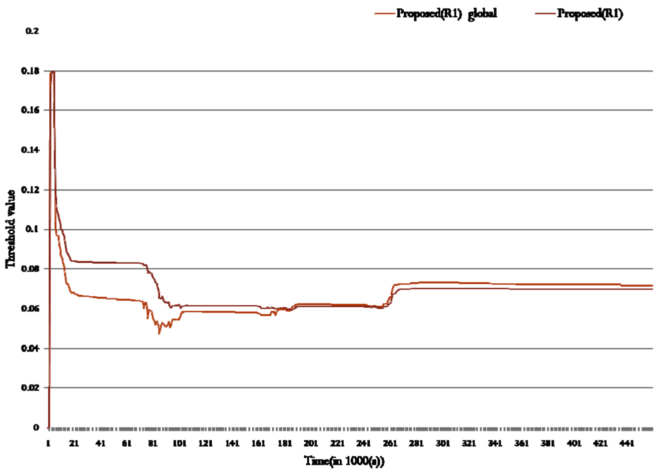


Fig. 10 The threshold for Cambridge trace

VI. Conclusions

In this paper, we propose a social-based data dissemination scheme for opportunistic networks. The proposed scheme combines three social metrics, transitivity, tie strength, and betweenness centrality, and uses threshold for the difference between total utility values.

For existing protocol, SimBetTS, we argue the problem of missing the indirect delivery probability. As ignoring the indirect delivery through intermediate nodes, it misses many chances to find better paths. To overcome this problem, we propose to incorporate transitivity, which considers the indirect delivery probability through the neighbor nodes, as one of forwarding metrics.

For removing the useless forwards, we suggest using threshold when forwarding so that the proposed scheme can avoid useless forwards when both nodes have few difference of total utility value. Additionally, in order to prevent messages from being isolated, we conditionally use betweenness centrality.

We implement an event-driven simulator which runs on real-world trace data sets, and compare the proposed scheme with the existing protocols, epidemic routing and SimBetTS. As mentioned before, epidemic routing gives the best performance of delivery ratio and delay time, but leaves the problem of overhead. Compared with SimBeTS, the proposed scheme achieves higher delivery ratio and shorter delay time while incurring less overhead.

As using the threshold for difference of total utility values, we reduce the useless forwards. Although we derive the value based on the average difference of total utility values per hop for all delivered messages to destination nodes, it is not fine-grained. As a future work, we are now trying to figure out how to select the threshold according to the context of networks, such as network size, number of replicas, etc. All experiments in this paper assume that all nodes have unlimited resource usages, such as buffer size and power consumption. But when the resources are limited, it is inevitable to fall down the performance. To recover the weakened performance by limited resource capacity is another topic for future research.

Acknowledgment

This work was supported at the IT R&D program of MKE/KEIT under grant KI001877 [Locational/Societal Relation-Aware Social Media Service Technology].

References

[1] Luciana Pelusi, Andrea Passarella and Marco Conti "Opportunistic networking: data forwarding in disconnected mobile ad hoc networks", IEEE Communications Magazine, November, 2006.

[2] Vahdat A. and Becker D., "Epidemic routing for partially connected ad hoc networks", Technical Report CS-2000-06, Computer Science Department, Duke University, 2000.

[3] Chiara Boldrini, et al., "Exploiting users' social relations to forward data in opportunistic networks: The HiBOp solution", in Pervasive and Mobile Computing, 2008.

[4] Elizabeth M. Daly and Mads Haahr, "Social network analysis for information flow in disconnected delay-tolerant MANETs", IEEE Transactions on Mobile Computing, Vol. 8, No. 5, May 2009.

[5] Hui P. and Crowcroft J, "Bubble Rap: Social-based forwarding in delay tolerant networks", International Symposium on Mobile Ad Hoc Networking & Computing (MobiHoc'08), May 2008.
[6] Boldrini, C., Conti, M., Iacopini, I. and Passarella, A. "Hi-BOp: a history based routing protocol for opportunistic networks", Paper presented in the Proceedings of the IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2007), pp.1–12.
[7] Elizabeth Daly and Mads Haahr, "Social network analysis for routing in disconnected delay-tolerant MANETs", MobiHoc'07, Sep. 9-14, 2007.
[8] Hui, P. and Crowcroft, J. "Bubble rap: forwarding in small world dtns in every decreasing circles", Technical report, Technical Report UCAM-CL-TR684. Cambridge, UK: University of Cambridge, 2007.
[9] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot,"Pocket Switched Networks and Human Mobility in Conference Environments," Proc. ACM SIGCOMM Workshop Delay-Tolerant Networking (WDTN '05), pp. 244-251, Aug. 2005.
[10] S.Boccaletti, et al., "Complex networks: structure and dynamics", Physics Reports 424, 175-308, 2006.
[11] M.E.J.Newman, "The Structure and function of complex networks", SIAM Review 45, 167-256, 2003.
[12] S. Milgram, "The Small World Problem," Psychology Today, vol. 2, pp. 60-67, May 1967.
[13] M.S. Granovetter, "The Strength of Weak Ties," The Am. J.Sociology, vol. 78, no. 6, pp. 1360-1380, May 1973.
[14] Hugh Louch, "Personal network integration: transitivity and homophily in strong-tie relations", Social Networks 22, 45-64, 2000.
[15] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms," Proc. IEEE INFOCOM, 2006.
[16] Duncan J. Watts, "Small worlds: the dynamics of networks between order and randomness", Princeton studies in complexity, 2003.
[17] Lindgren A., Doria A. and Schelen O., "Probabilistic routing in intermittently connected networks", ACM Mobile computing and Communications review, Vol. 7, pp.19-20, 2003.
[18] Burns, B., Brock, O. and Levine, B."MV routing and capacity building in disruption tolerant networks", Paper resented in the Proceedings of the 24th IEEE Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005).
[19] Burgess, J., Gallagher, B., Jensen, D. and Levine, B. "MaxProp: routing for vehicle-based disruption-tolerant networks", Paper presented in the Proceedings of the 25th IEEE Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2006).
[20] Spyropoulos, T., Psounis, K. and Ragavendra, C. "Efficient routing in intermittently connected mobile networks: the multiple-copy case", ACM/IEEE Transactions on Networking,Vol. 16, pp.77–90, 2007.
[21] Musolesi, M., Hailes, S. and Mascolo, C. "Adaptive routing for intermittently connected mobile ad hoc networks", Paper presented in the Proceedings of the IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2005), pp.183–189.
[22] Christian Kretschmer, et al.,"Delay-tolerant on-demand routing for mobile ad hoc networks", Computer science technical report No. 244, University of Freiburg, Germany, December 2008.
[23] Widmer, J. and Le Boudec, J. (2005) "Network coding for efficient communication in extreme networks", Paper presented in the Proceedings of the ACM SIGCOMM 2005 Workshop on Delay Tolerant Networking (WDTN 2005), pp.284–291