Incentive-Aware Data Dissemination in Delay-Tolerant Mobile Networks

Ting Ning, Zhipeng Yang, Xiaojuan Xie and Hongyi Wu

***Abstract*—This work centers on data dissemination in delay-tolerant mobile networks, where data fall into a range of interest types and each node may have one or multiple interests. The goal is to deliver data messages from sources to nodes with corresponding interests. We consider selfish nodes with rational behavior, and propose a credit-based incentive scheme to promote nodal collaboration. The key challenge is to effectively track the value of a message under such a unique network setting with intermittent connectivity and multiple interest types. Given poor end-to-end connections, credits are rewarded to the final deliverer only. Thus the value of a message for an intermediate node highly depends on its probability to deliver the message. Such probability itself is nontrivial to estimate. Moreover, a message is usually desired by multiple mobile users. Therefore, it can be potentially "sold" multiple times to different receivers. On the other hand, while more than one copies can be created during the transmissions of a message, a particular receiver "pays" for the first received copy only. These characteristics together make the development of incentive mechanism a unique, interesting, and challenging problem. In this paper, we present effective schemes to estimate the expected credit reward, and formulate nodal communication as a two-person cooperative game, whose solution is found by using the Nash Theorem. Extensive simulations are carried out based on real-world traces to evaluate the proposed scheme in terms of data delivery rate, delay and overhead. To our best knowledge, this is the first work that incorporates incentive stimulation into data dissemination in delay-tolerant mobile networks with selfish nodes and multiple interest types.**

I. INTRODUCTION

This work centers on data dissemination in a mobile wireless network that consists of a diversity of portable devices, e.g., cellular phones, PDAs, and laptops. In contrast to data communication via cellular channels that incur extra cost to users, the proposed scheme exploits free, low-power, short-range radio (e.g., Wifi or Bluetooth) commonly available on phones, PDAs, and laptops to establish an intermittently connected mobile network for data transmission. Due to the short radio communication range and unrestrained nodal mobility, the connectivity of the network is very low and dynamic, where a node connects to other nodes only occasionally, forming a delay-tolerant network (DTN) [1].

The data to be disseminated fall into a range of interest types, such as weather forecast, event alerts, commercial advertisement, movie trailers, blog updates, and various news. Such data are generated by their sources and accessible by mobile nodes via and access points (APs). A mobile node may wish to receive data in one or multiple interest types. When a node has access to an AP, it can download messages of its own and/or others' interests. When two nodes meet, they may communicate and share part or all of their data if an agreement is achieved.

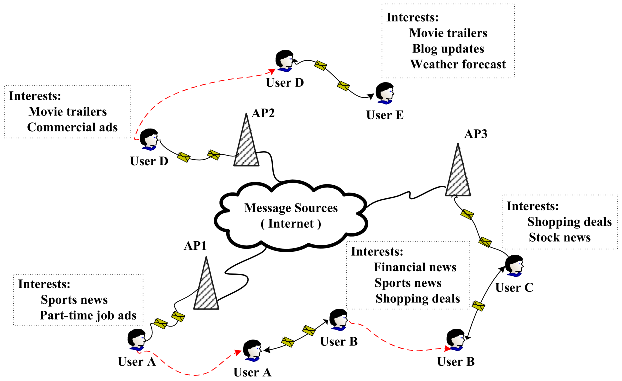


Fig. 1. An example of data dissemination.

An example is illustrated in Fig. 1, where Mobile User A is equipped with a smart phone and interested in sport news and part-time job advertisement. He can of course download such information via cellular channels, which is however costly. Thus, he intends to exploit the freely accessible APs scattered at such spots as coffee shops, restaurants, and campus buildings that he visits routinely. He can download data of sport news and part-time job advertisement to not only satisfy himself but also share them via bluetooth or Wi-Fi to other mobile users (e.g., User B) who do not have routine access to such free APs. In this way, data can be spread without specifying the recipients but only by declaring the type of the message content. Routing is no longer to find an end-to-end path. Instead, the network delivers data to the interested users based on the types of messages. Note that, a mobile user can intentionally download data that are not of his interests but desired by other users to further improve the efficiency of the entire network.

However, the participants in such a network can be either cooperative or selfish. If all nodes are cooperative, each of them carries messages for others voluntarily. On the other hand, if a node is selfish, it may be reluctant to consume its energy, buffer and bandwidth resources for other nodes, and thus refuse to carry any messages other than the ones interested by itself. In the worst case where every node is selfish, data are not shared at all among mobile nodes, leading to poor network performance. To this end, an incentive scheme is imperative to stimulate nodal cooperation.

In this work, we consider selfish nodes with rational be-

The authors are with The Center for Advanced Computer Studies, University of Louisiana at Lafayette, P.O. BOX 44330, Lafayette, LA 70504. This work is supported in part by the National Science Foundation under grant CNS-0831823.

havior. More specifically, a node is driven by its interests. It performs data transmission only if it gains benefit. Neither does it consume its resources to help nor to maliciously attack other nodes. Moreover, we assume authentication service is available so that a node can not cheat by intentionally forging its own identity to obtain free forwarding service or get more benefit from other nodes. The authentication mechanism is out of the scope of this paper. We propose a credit-based incentive scheme to promote nodal collaboration. The key challenge is to effectively track the value of a message under such a unique network setting with intermittent connectivity and multiple interest types. Given poor end-to-end connections, credits are rewarded to the final deliverer only. Thus the value of a message for an intermediate node highly depends on its probability to deliver the message. Such probability itself is nontrivial to estimate. Moreover, a message is usually desired by multiple users. Therefore, it can be potentially "sold" multiple times to different receivers. Meanwhile, while more than one copies can be created during the transmissions of a message, a particular receiver "pays" for the first received copy only. These characteristics together make the development of incentive mechanism a unique, interesting, and challenging problem. We present effective schemes to estimate the expected credit reward, and formulate nodal communication as a two-person cooperative game, whose solution is found by using the Nash Theorem. Extensive simulations are carried out based on real-world traces to evaluate the proposed scheme in terms of data delivery rate, delay and overhead. To our best knowledge, this is the first work that incorporates incentive stimulation into data dissemination in delay-tolerant mobile networks with selfish nodes and multiple interest types.

The rest of the paper is organized as follows. Sec. II discusses related work for cooperation stimulation and data transmission in delay-tolerant mobile networks. Sec. III introduces our proposed incentive scheme. Sec. IV presents simulation results. Finally, Sec. V concludes the paper.

II. RELATED WORK

In this section, we discuss cooperation stimulation in wireless networks and DTN data transmission protocols, which are highly related to this work.

A. *Incentive Approaches in Wireless Networks*

The problem of cooperation has been extensively studied in recent years, in the context of mobile ad-hoc networks, wireless mesh networks and peer-to-peer applications. In general, the incentive schemes fall into three categories: reputation-based, credit-based, and barter-based approaches.

Under reputation-based approaches [2], each node is associated with a reputation that reflects its degree of cooperation. Nodes build up their reputation by relaying packets for other nodes and are rewarded with high priority when their own packets are delivered across the network. The routing path is selected according to nodal reputation.

The credit-based scheme was first introduced in [3], where a node earns credits, i.e., the virtual money, by forwarding packets for others, and uses such credits to obtain forwarding services from other nodes in the network. Motivated by the concept of credit, several schemes are proposed to stimulate cooperation in packet forwarding [4]–[6].

In barter-based approaches, every node forwards as much traffic for a neighbor as the neighbor forwards for it. In other words, message forwarding is based on pair-wise exchange among nodes [7], [8].

However, none of the above incentive approaches are directly applicable in DTNs. Due to frequent partitioning and lack of an end-to-end contemporaneous path, it is impractical for a node to manage reputation of its neighbors, as required in the reputation-based approaches. It is also difficult to estimate the number of intermediate nodes that would participate in packet forwarding in order to set a proper initial credit for the packet, as required in the credit-based systems. In barter-based approaches, a reliable third party is imperative to monitor nodal behavior, such that a node can lower its service to a neighbor when the latter is detected with misbehaving. In DTNs, however, there are no such reliable third party nodes for behavior monitoring.

Besides the above stimulation mechanisms for well connected wireless networks, there are several incentive schemes recently proposed for DTNs. For example, the pair-wise Tit-for-Tat (TFT) is proposed in [9] as an incentive mechanism for DTNs, aiming to maximize nodal throughput under the constraints of TFT. In [10], every packet is encrypted, and nodes are stimulated for its own benefit to relay for other nodes in the network. Relay nodes are asked to show successful relay proofs to the giving nodes. Otherwise, the giving nodes broadcast misbehavior messages to the whole network, which will eventually exclude such misbehavior nodes from the network. [9] and [10] focus on unicast only in contrast to the one-to-many communication paradigm considered in our work. In [11], a credit-based incentive system is proposed for disruption tolerant networks, which integrates credit and cryptographic technique to solve the edge insertion and edge hiding attacks among nodes. It requires the source and destination always have access to a mostly available network for control message and a trusted third party is needed to manage the verification and payment services. In addition, [12] proposes socially selfish routing in DTNs, where a node exploits social willingness to determine whether or not to relay packets for others. These DTN incentive schemes consider different networking and application settings, and thus not readily applicable in this work.

B. *DTN Routing*

In DTNs, a routing decision is made in the absence of end-to-end paths and without the overall knowledge of the network. The earliest DTN routing protocols, e.g., epidemic routing [13], are based on controlled flooding. To reduce overhead, utility-based replication schemes are proposed [14]–[16], where a node estimates its probability to deliver data to the destination and replicates data over the best carriers according to utility metrics. Such metrics are largely derived from the temporal and/or spatial information based on nodal contacts.

Several algorithms are proposed recently for routing in social networks. For example, [17] employs two metrics, social similarity and betweenness, to estimate the importance and activities of the nodes in the network, and leverages such information to make decisions in message forwarding. [18] exploits the fact that people in the same community meet each other more frequently. It thus uses small community labels, and forwards data to the nodes that belong to the same community as the destination node, in order to achieve high delivery probability. [19] addresses publish-subscribe routing problem in human networks based on co-location and movement pattern of the hosts in the network. [20] models DTN routing as a utility-driven resource allocation problem and estimates per-packet utility to determine whether to replicate the packet.

However, all of these solutions assume that nodes in the network are cooperative and never decline service to other nodes. Such assumption is not always true in a real social network since the resources of a node, such as power, storage, and connection opportunities, are limited. As a result, a node might not be willing to provide free service to others unless they are rewarded. Therefore, the provision of incentive is imperative to stimulate the cooperation among the nodes.

III. THE PROPOSED INCENTIVE SCHEME

Under the unique setting of a DTN with selfish nodes and multiple interests types, we propose a novel incentive scheme that incorporates credit-based stimulation mechanism and interest-oriented delivery probability.

A. Definitions

To facilitate our exposition, we first introduce several basic definitions that serve as the basis for the proposed incentive scheme.

Definition 1. An *interest* is the type of data that a node wishes to acquire.

The examples of interests include blog updates, serial publication in installments, weather forecast, event alerts, commercial advertisement, and various news.

Definition 2. A *source of an interest* is a node that generates data messages that match the corresponding interest.

Definition 3. A *sink of an interest* is a node that wishes to acquire and consume data messages that match the corresponding interest.

A node may serve as the sources or sinks for multiple interests. At the same time, the data messages of a particular interest may be supplied by multiple sources, and multiple nodes may become the sinks of the same interest. Assume there are N interest types in the network. Each node serves as the sink of at least one and up to N interests.

Definition 4. The *effective interest contact probability (EICP)* of Node n in Interest i represents the likelihood that Node n contacts a sink of Interest i directly or indirectly.

The EICP indicates the probability that a node delivers data in an interest type to its sink(s). Its value intrinsically depends on nodal mobility, and aggregates direct and indirect contacts. The former, i.e., the *direct* interest contact probability of Node n in Interest i, indicates the probability that Node i directly meets a node that is the sink of Interest i. The latter is the *indirect* interest contact probability, which captures the probability that Node i delivers data to the sinks of Interest i via other nodes indirectly. In this research, we adopt the exponentially weighted moving average (EWMA), which is one of the most effective schemes for online estimation and has been employed in many applications, to maintain and update the nodal effective interest contact probability [21]. EWMA is simple, needs low computation, reacts to small shifts timely and requires only constant storage space. When two nodes meet each other, any one of them can perform the computation and update the nodal effective interest contact probability.

More specifically, each node maintains a timer. If there is no contact with other nodes within an interval of Δ, the timer expires, generating a timeout event. Let $\vartheta_n(i)$ denote the direct contact probability of Node n in Interest i. $\vartheta_n(i)$ is initialized to zero, and updated at every contact with the node matching Interest i or a timeout event, whichever comes first. The update function is given below.

$$\vartheta_n(i) = \begin{cases} (1-\alpha)\vartheta_n(i) + \alpha & \text{Contact} \\ (1-\alpha)\vartheta_n(i) & \text{Timeout,} \end{cases} \tag{1}$$

where $0 \leq \alpha \leq 1$ is a constant employed to keep partial memory of historic contact status.

Let $\xi_n(i)$ denote the indirect contact probability of Node n in Interest i. It is updated as follows:

$$\xi_n(i) = \begin{cases} (1-\beta)\xi_n(i) + \beta\vartheta_k(i) & \text{Contact} \\ (1-\beta)\xi_n(i) & \text{Timeout,} \end{cases} \tag{2}$$

where $0 \leq \beta \leq 1$ is the historic factor and $\vartheta_k(i)$ is the direct contact probability to any nodes with Interest i by Node k (which Node n meets).

Let $\chi_n(i)$ denote the EICP of Node n in Interest i. We assume the direct and indirect contacts are independent and integrate them to obtain EICP:

$$\chi_n(i) = 1 - (1 - \vartheta_n(i))(1 - \xi_n(i)), \tag{3}$$

where $\vartheta_n(i)$ and $\xi_n(i)$ are the direct and indirect contact probability of Node n in Interest i, respectively.

Definition 5. The *credit* is the virtual currency employed for stimulating cooperation among selfish nodes.

Let λ_n denote the amount of credits Node n owns. λ_n is initialized to Λ and updated during messages transmissions to be discussed in Sec. III-B.

Definition 6. The *rewarding policy* defines how a node gains credits.

To support the two-person cooperative game to be discussed next in Sec. III-C, we adopt a simple rewarding policy as follows. If Node n receives a message that matches its interests from Node m, the former rewards one credit to the latter. In a transmission where the data message does not match the receiver's interest, the nodes do not gain or pay any credits.

Definition 7. Each data message is associated with a *duplication degree*, which indicates the number of copies it has.

Let $C^m(i)$ denote the estimated number of copies of Message m in Interest i. Split-based approach [22] is adopted to estimate $C^m(i)$. More specifically, each data message maintains a parameter, called duplication degree, that is an indicator of the current estimation of copies of the message in the network. Let $S_n^m(i)$ denote the duplication degree of Message m in Interest i in Node n. The parameter is initialized to 1 by the source. When Nodes n and k meet, they update their current parameters as $S_n^m(i) = S_k^m(i) = (S_n^m(i) + S_k^m(i))/2$, if they both have a copy of the message. The same process applies when Node n receives a new message from Node k, where $S_n^m(i)$ is 0 before the calculation. While dynamics exist before the splitting process converges, $C^m(i) = 1/S_n^m(i)$ or $1/S_k^m(i)$ serves as a good estimation of the number of copies of Message m in the network in general.

Definition 8. Each data message is associated with a *message appraisal*, which indicates its potential value.

Let $A^m(i)$ denote the appraisal of Message m in Interest i. Based on the above rewarding policy, $A^m(i)$ is measured by the number of sinks of Interest i that have not received Message m, i.e., the potential credits that a node can gain by delivering the message to those sinks. The appraisal of a message is set by its source at the time of generation. In this work, we assume the sources generate messages and they set the appraisal value by the total number of sinks of the corresponding interest, which can be estimated through the subscription process.

Definition 9. The *expected credit reward* is the anticipated earning for a node to obtain a data message.

The expected credit reward is employed to facilitate a node to determine if a message should be traded. Let $R_n^m(i)$ denote the expected credit reward if Node n trades in Message m of Interest type i. $R_n^m(i)$ is clearly dependent on the message's appraisal (i.e., $A^m(i)$), which indicates its potential value. Moreover, it depends on two other factors. First, since Node n only gains credits when it delivers the message to corresponding sinks, $R_n^m(i)$ depends on the delivery probability of Node n. Intuitively, the direct interest contact probability (i.e., $\vartheta_n(i)$) should be considered here, as only direct delivery to a sink gains a credit under our rewarding policy. However, we propose to employ $\chi_n(i)$, in order to reflect the anticipated reward due to the credits gained via not only direct delivery of Message m but also the delivery of other message(s) traded by Message m. Second, a sink is interested in one copy of a message and thus pays for the first received copy only. However, most practical DTN routing protocols create multiple copies for each message. The more the copies, the lower the probability for a copy to be delivered to a sink before other copies. Thus we have $R_n^m(i)$ defined below:

$$R_n^m(i) = A^m(i) * \chi_n(i)/C^m(i). \tag{4}$$

Definition 10. When Node n meets another node, it needs to decide whether or not to exchange messages with the latter. Due to its selfish nature, Node n aims to maximize its expected credit reward should message exchange happen. The utility function used by Node n in making such a decision is as below:

$$Max \quad U_n = \sum_{i=1}^{I}\Big(\sum_{m\in\phi(i)} R_n^m(i) - \sum_{m\in\varphi(i)} R_n^m(i)\Big), \tag{5}$$

where U_n is the utility function of Node n, I is the total number of interest types, $\phi(i)$ and $\varphi(i)$ are set of messages in Interest i after and before exchange, respectively.

B. Overview of The Proposed Incentive Scheme

Based on above definitions, our proposed incentive scheme is outlined below. To facilitate our discussions, we assume that each data message, e.g., Message m at Node n, is associated with a descriptive metadata, which includes its interest type (i.e., i), sequence number (i.e., m), appraisal (i.e., $A^m(i)$), and duplication degree (i.e., $S_n^m(i)$). Let $\hat{L}_n$ denote the set of metadata of the messages at Node n, and Φ_n the list of EICPs of Node n (i.e., $\Phi_n = \{\xi_n(i) | 1 \le i \le I\}$).

When Node n meets another node, e.g., Node k, it follows five steps to acquire necessary information and make decisions on trade of messages.

Step (1) Node n sends to Node k a control message, which includes $\hat{L}_n$ and Φ_n. Meanwhile, Node n receives a similar control message from Node k too. Both nodes update their EICPs according to Eqs. (1)-(3).

Step (2) Node n creates its candidate list $\tilde{L}_n$, that includes messages available at Node k but not at Node n, i.e., $\tilde{L}_n = \hat{L}_k - (\hat{L}_n \cap \hat{L}_k)$. $\tilde{L}_n$ also excludes any messages that Node n has received before. Similarly, Node k creates its candidate list $\tilde{L}_k$.

Step (3) Node n checks if it is a sink for any messages in $\tilde{L}_n$. Let $\tilde{P}_n$ denote the set of such messages and p_n=$min(\lambda_n, |\tilde{P}_n|)$. Node n pays Node k one credit for transmitting a message in $\tilde{P}_n$ from the latter to the former, up to p_n messages. Accordingly, Node n's candidate list removes those messages it has just received, and updates its available credit as $\lambda_n = \lambda_n - p_n$. Similar actions are taken by Node k.

Step (4) Nodes n and k then bargain which messages should be traded. This process is formulated as a two-person cooperative game and Nash Theorem is applied to reach optimal solution, yielding two final message lists that Nodes n and k will exchange, denoted by L_n and L_k with $|L_n| = |L_k|$. The two-person cooperative game model and its solution based on Nash theory will be elaborated next in Sec. III-C.

Step (5) Finally, Nodes n and k trade messages, pair by pair, in L_n and L_k.

C. Game Theory Model and Nash Solution

While most steps of the proposed incentive scheme are straightforward, Step 4 is based on the two-person cooperative game theory and worth further elaboration.

By assuming that the nodes are rational (i.e., their actions are driven by their interests), the interaction between two nodes can be modeled as a two-person cooperative game, which was proposed by John Forbes Nash in [23]. The two persons in the game are rational and selfish. In other words, the player does not act maliciously to hurt the other's interest, providing the foundation for cooperation. On the other hand, the selfishness naturally leads to different objective for each person. Note that such selfish behavior may promote or impede cooperation, depending on whether it brings benefit or harm to the players. For example, under our network setting, the objective of a selfish node is to maximize its utility function defined in Eq. (5). A node can usually improve its utility function by obtaining a new data message. This, however, increases the duplication degree of the data message and thus decreases the utility function of the other node. The two person cooperative game allows players to reach a binding agreement, based on their possibly conflicting interests. This is in a sharp contrast to non-cooperative game that disallows binding agreement between players. The selfishness of the players ensures that, if a binding agreement is reached, it must benefit both players.

The optimal solution (or Nash solution) for two-person cooperative game is given by maximizing the Nash product [23], i.e.,

$$(\hat{U}_n, \hat{U}_k) = arg\ \ max(U_n - D_n) \times (U_k - D_k), \qquad (6)$$

where (D_n, D_k) is the status quo point, which is set to (0,0) in this work. $\hat{U}_n$ and $\hat{U}_k$ are the utility gains of Nodes n and k in the Nash solution, respectively, which must be positive to arrive at a feasible solution. Note that U_n (or U_k) is determined by the set of messages exchanged as shown in Eq. (5). Therefore, the optimal solution yields an optimal set of messages that should be exchanged between Nodes n and k.

Although the general optimal solution is defined above, it is nontrivial to find such a solution under a practical network setting, because each node may own a large number of data messages and it takes exponential time to identify $\hat{U}_n$ and $\hat{U}_k$ among all possible solutions. To this end, we adopt a simple heuristic approach by considering one pair of messages at a time. More specifically, for each pair of messages, $m_n \in L_n$ and $m_k \in L_k$, the corresponding Nash product is calculated by by assuming Messages m_n and m_k were exchanged between Nodes n and k. The pair that results in the maximum Nash product is selected for trade. The time complexity is $|L_n| \times |L_k|$. The above process repeats until Nodes n and k lose their connection, or L_n or L_k is empty, or there is no solution for Eq. (6).

Note that, when Node n sends Message m to Node k, it still keeps a copy of the message, but with an updated duplication degree. Therefore, the data queue of a node is likely full at the steady state of the network. Receiving a message means to replace an existing message with the lowest expected credit reward. The loss of such a message is taken into account in the calculation of utility gain, i.e., the dropped message is excluded from $\phi(i)$ in Eq. (5).

TABLE I
OVERALL PERFORMANCE COMPARISON BASED ON HAGGLE TRACE.

	Data Delivery Rate	Delay	Overhead
Direct	0.42	36109s (10.1h)	1
SelfExchange	0.58	22510s (6.25h)	1
CooperRdm	0.67	27653s (7.68h)	34
Incentive	0.82	10238s (2.84h)	2
Cooperative	0.86	8764s (2.43h)	10

TABLE II
OVERALL PERFORMANCE COMPARISON BASED ON DIESELNET TRACE.

	Data Delivery Rate	Delay	Overhead
Direct	0.27	56850s (15.7h)	1
SelfExchange	0.48	34696s (9.6h)	1
CooperRdm	0.59	38615s (10.1h)	22
Incentive	0.70	13684s (3.8h)	3
Cooperative	0.76	11810s (3.3h)	9

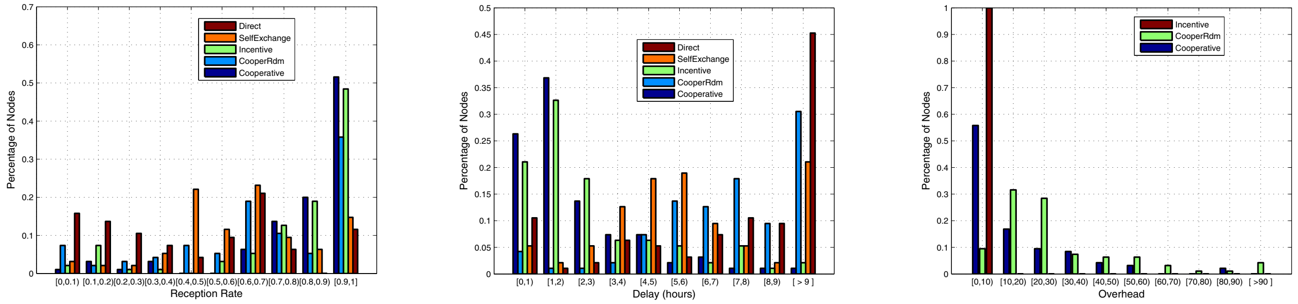
IV. SIMULATION RESULTS

We have carried out simulations to evaluate our proposed incentive mechanism. In this section, we first introduce our simulation setup, followed by performance comparison and discussions.

A. Simulation Setup

Due to the uniqueness of our networking setting, the mechanisms that described in Sec. II are not comparable with our scheme. We compare different schemes by varying the degree of cooperation of nodes: the "Direct" scheme, where no cooperation exists among nodes and thus a node only downloads its interested messages from corresponding APs; the "SelfExchange" scheme, where a node can obtain its interested messages from an AP or a mobile node but does not carry any messages out of its interests; the "CooperRdm" scheme, where the encountered nodes randomly choose some messages (besides their own interested messages) to exchange; the "Cooperative" scheme, where nodes are fully cooperative and always choose the most valuable messages to carry after satisfying its own interests; and our proposed incentive scheme (denoted by "Incentive").

Trace data of the Cambridge Haggle project [24] and UMass DieselNet project [25] are employed in our simulations, which represent human social networks and vehicular networks, respectively. In the Haggle project, mobile nodes called iMotes were distributed to 50 people attending IEEE INFOCOM workshop. Its datasets include contacts between iMotes and Bluetooth devices. We adopt the dataset which involves 98 iMotes and Bluetooth devices and run the simulation for a total period of $342,915$ seconds (or 3 days). In the UMass DieselNet project, a DTN testbed is constructed by $30-40$ transit buses, serving an area of approximately 150 square miles. The trace data provides contact events between buses.

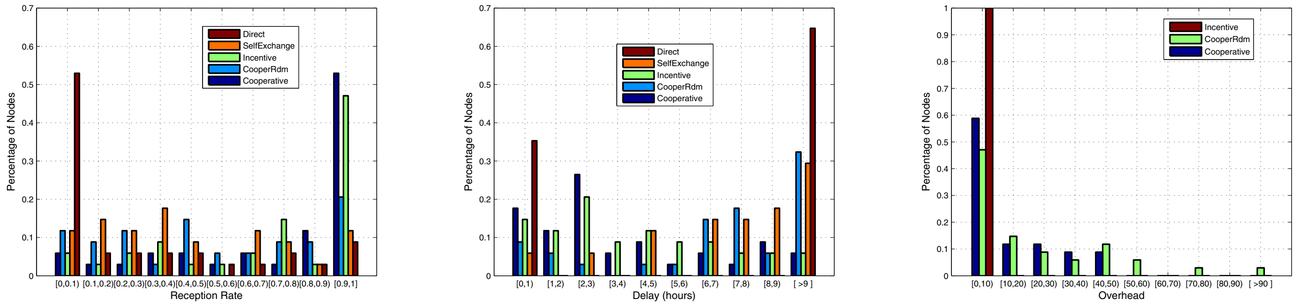


(a) Reception Rate Distribution.

(b) Delay Distribution.

(c) Overhead Distribution.

Fig. 2. Distribution of reception rate, delay and overhead under the Haggle Trace.



(a) Reception Rate Distribution.

(b) Delay Distribution.

(c) Overhead Distribution.

Fig. 3. Distribution of reception rate, delay and overhead under the DieselNet Trace.

Our simulation is based on the trace data obtained in 2006 with 37 buses [26] for a period of $1,250,847$ seconds (or about two weeks).

As the default setting of our simulations, we assume there are three APs and 15 interest types in the network. The message generation rate of the source is one message per 100 seconds. The queue size of each node is 100 messages. The values of α and β in Eqs. (1) and (2) are both 0.1, which are the best to reflect the impact of the historic status. We randomly choose some nodes as the APs and average the results over 10 simulation runs.

Our goal is to disseminate data messages to corresponding sinks with both delay and traffic overhead as low as possible. In our simulation studies, we are interested in the following metrics for performance evaluation: the network-wide data delivery rate, the distribution of nodal reception rates, the average delivery delay, and the message forwarding overhead. The network data delivery rate is defined as the ratio of the total number of delivered messages to the total number of messages that should be disseminated to the network. The reception rate of a node is the ratio of the total number of its interested messages received to the total number of such messages generated by the sources. The average delivery delay is a measure of how long a node waits to get an interested message. Message forwarding overhead is a cost factor, defined as the communication cost of a delivered message. Lower overhead means fewer transmission traffic in the network.

B. *Performance Comparison*

Tables I and II compare the overall performance of different schemes based on the Haggle trace and the DieselNet trace, respectively. They both show that the Cooperative scheme achieves the highest network data delivery rate, followed by Incentive, CooperRdm, SelfExchange and Direct schemes. The highest delivery rate of the Cooperative scheme is attributed to the fact that the nodes are altruistic to help each other by always choosing the most valuable messages to carry after satisfying its own interests, thus the chance of a message being delivered to the interested nodes therein is the highest among all schemes. At the same time, we notice that although the CooperRdm scheme is also fully cooperative and nodes are willing to carry others' messages for free, its delivery rate is much lower than the proposed Incentive scheme, serving as an evidence of the importance to exploit the estimated value of a message to make an appropriate decision on whether or not to carry it, especially under limited storage resources. In our proposed Incentive scheme, the estimated value of messages effectively fosters cooperation among nodes and makes efficient use of communication resource (that is determined by the capacity of nodes and their meeting opportunities), thus leading to higher delivery rate. Finally, if messages are simply forwarded at random, the chance for its delivery is low. Since the nodes in the SelfExchange scheme only exchange their interested messages and refuse to carry messages of other nodes' interests, the message dissemination is greatly hindered, resulting in a low delivery rate. Similarly, nodes under the Direct scheme do not cooperate at all, resulting in the lowest delivery rate.

The average delays of Cooperative and Incentive schemes are much shorter than other schemes, because both of them leverage the message value to estimate the probability to deliver the message and choose the best routes to forward it, thus delivering the message in a shorter time. Although the delivery rate and average delay of Cooperative scheme are better than the Incentive scheme, we can see that the former has a much higher overhead than the latter, because its altruism leads to more messages to be duplicated and distributed in the network. In a contrast, the proposed Incentive scheme achieves very low overhead, since a node receives a message copy only if it is confirmed that this copy can benefit it. Clearly, the overhead of Direct and SelfExchange is always one, because a node only receives its interested messages.

Figs. 2 and 3 illustrate the distribution of reception rate, average delay and overhead among nodes, under the Haggle trace and the DieselNet trace, respectively. The overhead of Direct and SelfExchange are omitted because they are always one as discussed above. Fig. 2 shows that about 48% of nodes under the proposed Incentive scheme can receive more than 90% of their interested messages, compared with 11% in Direct, 13% in SelfExchange, and 35% in CooperRdm. Although 52% of nodes in Cooperative achieve higher than 90% of reception rate, it is at the cost of 30% of the nodes with overhead greater than 20. This indicates that those nodes always play the role of relays and contribute much more than others to the network. In a sharp contrast, the proposed Incentive scheme stimulates the cooperation among nodes and allows a node to strike the balance between its individual interests and contribution to the network. As a result, none of the nodes have their overheads greater than 10. As shown in Fig. 2(b), about 56% of nodes in the Incentive scheme receive the interested message in less than 2 hours, while the average delay of more than 20% of nodes in Direct, SelfExchange and CooperRdm schemes is greater than 9 hours. Similar trend is observed in Fig. 3, i.e., the results under the DieselNet trace.

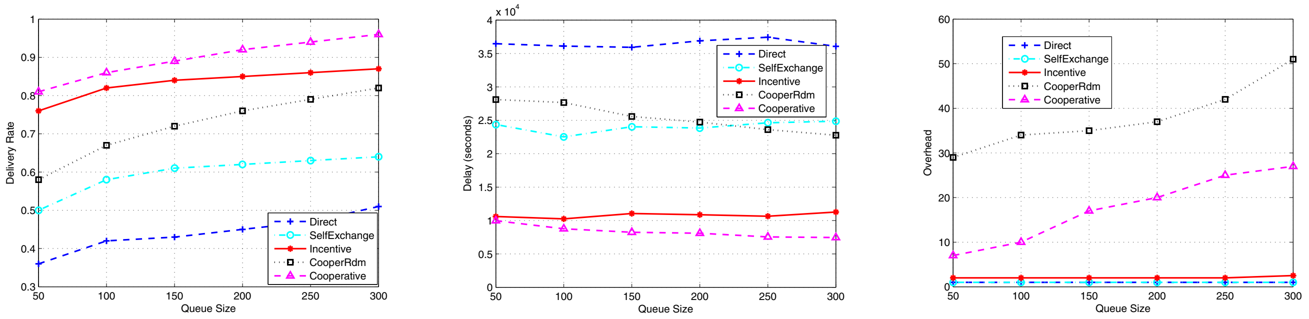
The impact of queue size is illustrated in Fig. 4. With the increase of the queue size, the data delivery rate of all schemes improves. Particularly, the delivery rate of Cooperative and CooperRdm increases significantly. This is because longer queue size allows them to buffer more data messages that can be exchanged and delivered later, thus increasing the delivery rate. At the same, their average delay decreases. The accompanying side effect is the rapidly increasing overhead. On the other hand, the increase of queue size has less impact on the performance of the proposed Incentive scheme, because it exchanges messages based on interests of individual nodes, aiming to promote rewards. Thus the nodes do not aggressively utilize the additional queuing space.

Fig. 5 shows the performance of different schemes by varying the number of APs. All APs are connected to the sources (e.g., via the Internet) that generate messages at a certain rate. Increasing the number of APs means that nodes can have more opportunities to access data sources and update their message buffers. This explains why the delivery rate of all schemes increases and the average delay decreases in Figs. 5(a) and 5(b). We also note that the average delivery rate of Cooperative is only about 3% better than the proposed Incentive scheme, while the overhead is about 5 times higher as shown in Fig. 5(c). It means that the exchange process in Incentive greatly reduces the transmission cost and at the same time maintains a preferable delivery rate and average delay.

Fig. 6 compares the performance by varying message generate rate. As shown in Fig. 6(a), with the increase of message generate rate, the delivery rate of Direct, SelfExchange and CooperRdm decreases, while the delivery rate of Incentive and Cooperative is stable. A higher generate rate means that more messages are generated by the sources. The more messages the sources generate, the higher workload the network must carry. If nodes don't contribute much to the network, such as in the Direct and SelfExchange, more messages would reside at sources and can't be disseminated to corresponding sinks in an acceptable delay time, thus resulting in decreasing delivery rate. In the Cooperative scheme, nodes are trying to distribute messages as many as possible and contribute their buffers and energy altruistically. That's the reason it can maintain the delivery rate stable. One might wonder why the delivery rate of the proposed Incentive scheme does not decrease. The reason lies in that fresh messages always have greater values and selfish nodes are willing to get those valuable ones in order to maximize their rewards. The cost of maintaining the delivery rate is the increasing overhead as shown in Fig. 6(c), since more messages are duplicated during their transmissions. With the increase of message generation rate, delay increases as demonstrated in Fig. 6(b). Given the limited resources at individual nodes, the higher the generate rate, the longer the messages need to reside at the sources and intermediate nodes, leading to longer average delay.

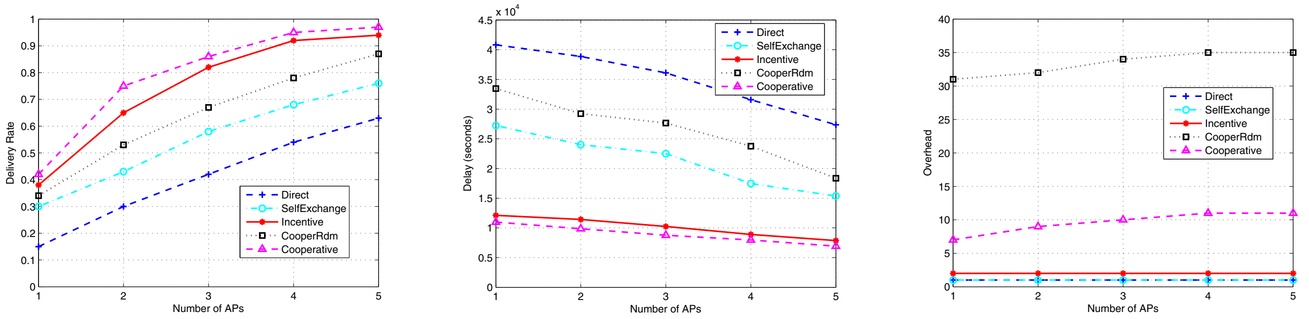
Fig. 7 depicts the impact of the number of interest types in the network. We assume each node can have at least one and up to N interests, where N is the max number of interest types. Fig. 7(a) shows that, with the increase of the number of interest types, the delivery rate decreases under SelfExchange, CooperRdm and Cooperative scheme, but increases slightly in the Incentive scheme. This is because cooperation between nodes in Incentive improves the possibility of delivering messages to each other when nodes have more interests. We observe in Fig. 7(b) that the average delay of all schemes slightly increases except the Direct scheme. From Fig. 7(c), we notice that the overhead of the proposed Incentive scheme is stable. This is because a node only chooses the most beneficial messages to trade and never wastes its resources to free-riders.

Fig. 8 illustrates the impact of the number of sinks per interest type in the network. Each interest type can be associated with multiple nodes. The more nodes in one interest type, the more overlapped interests two nodes may have. As the number of sinks per interest type increases, delivery rate increases



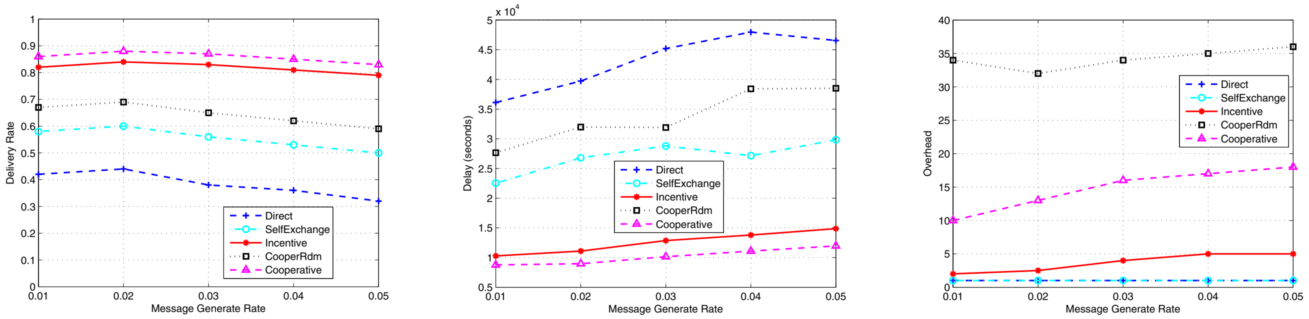
(a) Delivery Rate. (b) Delay. (c) Overhead.

Fig. 4. Variation of queue size.



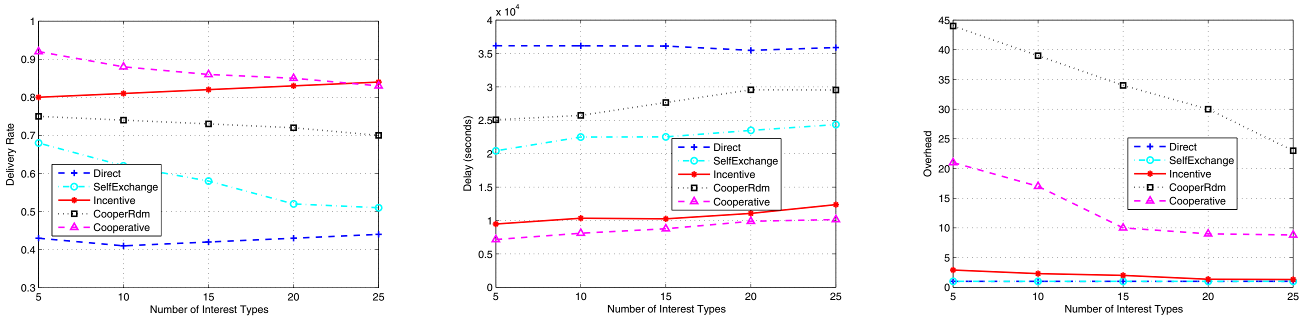
(a) Delivery Rate. (b) Delay. (c) Overhead.

Fig. 5. Variation of the number of APs.



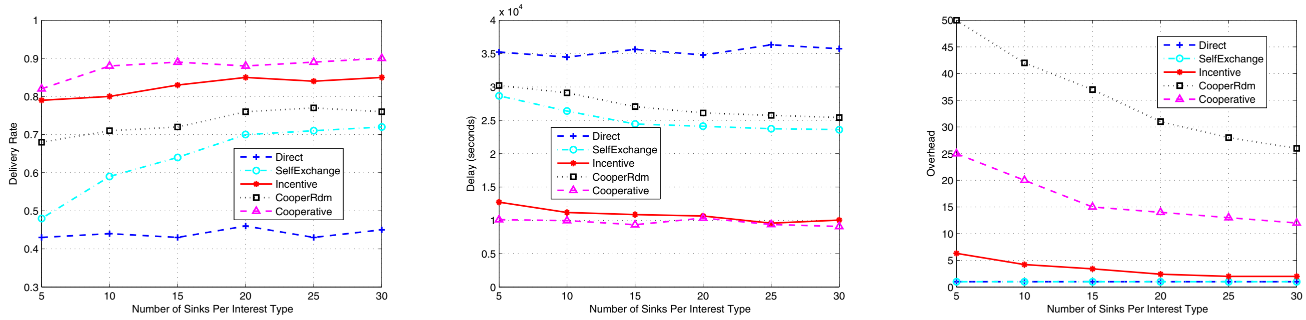
(a) Delivery Rate. (b) Delay. (c) Overhead.

Fig. 6. Variation of data message generate rate.



(a) Delivery Rate. (b) Delay. (c) Overhead.

Fig. 7. Variation of the number of interest types.



(a) Delivery Rate. (b) Delay. (c) Overhead.

Fig. 8. Variation of the number of sinks per interest type.

slightly, while the delay and overhead decrease, as illustrated in Figs. 8(a), 8(b) and 8(c), respectively. The reason is that with more common interests between nodes, there is a higher chance that a node can get its interested messages from others and thus the messages are disseminated in a shorter time.

V. CONCLUSION

In this work we have studied the problem of data dissemination in delay-tolerant mobile networks. We have considered selfish nodes with rational behavior, and proposed a credit-based incentive scheme to promote nodal collaboration. The key challenge is to effectively track the value of a message under such a unique network setting with intermittent connectivity and multiple interest types. These characteristics make the development of incentive mechanism a unique, interesting, and challenging problem. In this paper, we have presented effective schemes to estimate the expected credit reward, and formulated nodal communication as a two-person cooperative game, whose solution has been given by the Nash Theorem. Extensive simulations have been carried out based on real-world traces to evaluate the proposed scheme in terms of data delivery rate, delay and overhead. To our best knowledge, this is the first work that incorporates incentive stimulation into data dissemination in delay-tolerant mobile networks with selfish nodes and multiple interest types.

REFERENCES

[1] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets," in *Proc. of SIGCOMM*, pp. 27–34, 2003.

[2] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks," in *Proc. of MobiCom*, pp. 255–265, 2000.

[3] L. Buttyan and J. P. Hubaux, "Enforcing Service Availability in Mobile Ad-hoc WANs," in *Proc. of MoBiHoc*, pp. 87–96, 2000.

[4] S. Zhong, J. Chen, and Y. R. Yang, "Sprite, A Simple, Cheat-proof, Credit-based System For Mobile Ad-hoc Networks," in *Proc. of INFOCOM*, pp. 1987–1997, 2003.

[5] N. B. Salem, L. Buttyan, J. P. Hubaux, and M. Jakobsson, "A Charging and Rewarding Scheme for Packet Forwarding in Multi-hop Cellular Networks," in *Proc. of MoBiHoc*, pp. 13–24, 2003.

[6] M. Jakobsson, J. P. Hubaux, and L. Buttyan, "A Micropayment Scheme Encouraging Collaboration in Multi-hop Cellular Networks," in *Proc. of Financial Cryptography*, pp. 15–33, 2003.

[7] V. Srinivasan, P. Nuggehalli, and C. Chiasserini, "Cooperation in Wireless Ad Hoc Networks," in *Proc. of INFOCOM*, pp. 808–817, 2003.

[8] X. Xie, H. Chen, and H. Wu, "Bargain-based Stimulation Mechanism for Selfish Mobile Nodes in Participatory Sensing Network," in *Proc. of SECON*, pp. 72–80, 2009.

[9] U. Shevade, H. Song, L. Qiu, and Y. Zhang, "Incentive-aware Routing in DTNs," in *Proc. of ICNP*, pp. 238–247, October 2008.

[10] A. Mei and J. Stefa, "Give2Get: Forwarding in Social Mobile Wireless Networks of Selfish Individuals," in *Proc. of ICDCS*, pp. 488–497, 2010.

[11] B. Chen and M.C.Chan, "MobiCent: a Credit-Based Incentive System for Disruption Tolerant Network," in *Proc. of INFOCOM*, pp. 1–9, 2010.

[12] Q. Li, S. Zhu, and G. Cao, "Routing in Socially Selfish Delay Tolerant Networks," in *Proc. of INFOCOM*, pp. 857–865, 2010.

[13] A. Vahdat and D. Becker, "Epidemic Routing for Partially Connected Ad Hoc Networks," *Technical Report CS-200006, Duke University*, 2000.

[14] J. Leguay, T. Friedman, and V. Conan, "Evaluating Mobility Pattern Space Routing for DTNs," in *Proc. of INFOCOM*, pp. 276–283, 2006.

[15] J. Lebrun, C.-N. Chuah, D. Ghosal, and M. Zhang, "Knowledge-based Opportunistic Forwarding in Vehicular Wireless Ad Hoc Networks," in *IEEE Vehicular Technology Conference*, pp. 2289–2293, 2005.

[16] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic Routing in Intermittently Connected Networks," *SIGMOBILE Mobile Computing and Communication Review*, vol. 7, no. 3, pp. 1559–1662, 2003.

[17] E. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant Manets," in *Proc. of MoBiHoc*, pp. 32–40, 2007.

[18] P. Hui and J. Crowcroft, "How Small Labels Create Big Improvements," in *Proc. of the Fifth IEEE International Conference on Pervasive Computing and Communications Workshop*, pp. 65–70, 2007.

[19] P. Costa, C. Mascolo, M. Musolesi, and G. P. Picco, "Socially-Aware Routing for Publish-Subscribe in Delay-Tolerant Mobile Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 748–760, 2008.

[20] A. Balasubramanian, B. Levine, and A. Venkataramani, "DTN Routing As a Resource Allocation Problem," in *Proc. of ACM SIGCOMM*, pp. 373–384, 2007.

[21] Y. Wang and H. Wu, "DFT-MSN: The Delay Fault Tolerant Mobile Sensor Network for Pervasive Information Gathering," in *Proc. of INFOCOM*, pp. 1–12, 2006.

[22] M. Jelasity, A. Montresor, and O. Babaoglu, "Gossip-based Aggregation in Large Dynamic Networks," *ACM Transactions on Computer Systems*, vol. 23, no. 1, pp. 219–252, 2005.

[23] J. F. Nash, "Two-Person Cooperative Games," *Econometrica*, vol. 21, pp. 128–140, 1953.

[24] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace cambridge/haggle/imote/infocom (v. 2006-01-31)." Downloaded from http://crawdad.cs.dartmouth.edu/cambridge/haggle/imote/infocom, Jan. 2006.

[25] J. Burgess, B. N. Levine, R. Mahajan, J. Zahorjan, A. Balasubramanian, A. Venkataramani, Y. Zhou, B. Croft, N. Banerjee, M. Corner, and D. Towsley, "CRAWDAD data set umass/diesel (v. 2008-09-14)." Downloaded from http://crawdad.cs.dartmouth.edu/umass/diesel, Sept. 2008.

[26] J. Burgess, B. N. Levine, R. Mahajan, J. Zahorjan, A. Balasubramanian, A. Venkataramani, Y. Zhou, B. Croft, N. Banerjee, M. Corner, and D. Towsley, "CRAWDAD trace umass/diesel/transfer/spring2006 (v. 2007-12-02)." Downloaded from http://crawdad.cs.dartmouth.edu/umass/diesel/transfer/spring2006, 2007.