

Cellular Traffic Offloading through Opportunistic Communications Based on Human Mobility

Zhigang Li¹, Yan Shi¹, Shanzhi Chen^{2,3} and Jingwen Zhao⁴

¹ State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876 - China

[e-mail: zhigang.li.mail@gmail.com, shiyan@bupt.edu.cn]

² State Key Laboratory of Wireless Mobile Communications, China Academy of Telecommunications Technology, Beijing 100083 – China

³ Datang Telecom Technology & Industry Group
Beijing 100083 – China

[e-mail: chensz@datanggroup.cn]

⁴ International school, Beijing University of Posts and Telecommunications,
Beijing 100876 - China

[e-mail: 2011213263@bupt.edu.cn]

*Corresponding author: Yan Shi

*Received December 11, 2014; revised February 7, 2015; accepted February 27, 2015;
published March 31, 2015*

Abstract

The rapid increase of smart mobile devices and mobile applications has led to explosive growth of data traffic in cellular network. Offloading data traffic becomes one of the most urgent technical problems. Recent work has proposed to exploit opportunistic communications to offload cellular traffic for mobile data dissemination services, especially for accepting large delayed data. The basic idea is to deliver the data to only part of subscribers (called target-nodes) via the cellular network, and allow target-nodes to disseminate the data through opportunistic communications. Human mobility shows temporal and spatial characteristics and predictability, which can be used as effective guidance efficient opportunistic communication. Therefore, based on the regularity of human mobility we propose NodeRank algorithm which uses the encounter characteristics between nodes to choose target nodes. Different from the existing work which only using encounter frequency, NodeRank algorithm combined the contact time and inter-contact time meanwhile to ensure integrity and availability of message delivery. The simulation results based on real-world mobility traces show the performance advantages of NodeRank in offloading efficiency and network redundant copies.

Keywords: Cellular traffic offloading, human mobility, opportunistic communications

This work is partially funded by the National Science Foundation Projects (Grant No. 61300183) and the National Science Fund for Distinguished Young Scholars (Grant No. 61425012) in China.

1. Introduction

In the past several years, the explosive increase of intelligent mobile devices such as smartphones, tablets, and laptops, especially the increasing popularity of various applications for smartphones has led to the rapid growth of data traffic in cellular network. Global mobile data traffic will increase nearly 11-fold between 2013 and 2018. Mobile data traffic will grow at a compound annual growth rate (CAGR) of 61 percent from 2013 to 2018, reaching 15.9 exabytes per month by 2018 [1]. The unpredictable continuous growth of data traffic will bring great challenge to the current cellular network. How to reduce the congestion of network, improve the quality of service for customers and reduce the cost become the most urgent problems to solve at present.

The mobile data dissemination in mobile networks may include multimedia newspapers, weather forecasts, movie trailers, mobile social applications, shopping advertisements, etc. All these information is generated by content service providers. If everyone gets information from cellular network infrastructure like Fig. 1(a), the load of the cellular network infrastructure will be high. In order to save the bandwidth of cellular network infrastructure, benefit from the delay-tolerant nature of non-real-time applications, these service providers through the base station may deliver information to only a small fraction of selected users (i.e. target-nodes, like the N_1, N_2, N_3 in Fig. 1(b)), then the target-nodes can further disseminate the information to subscribed users when their mobile phones are in proximity and can communicate opportunistically using Wi-Fi or Bluetooth technology as shown in Fig. 1(b) to reduce cellular data traffic. The service provider will send the information to N_4 directly through cellular networks, if he or she fails to receive the information before their delivery deadline. We can translate this objective into maximizing the expected number of users that can receive the delivered information through opportunistic communications.

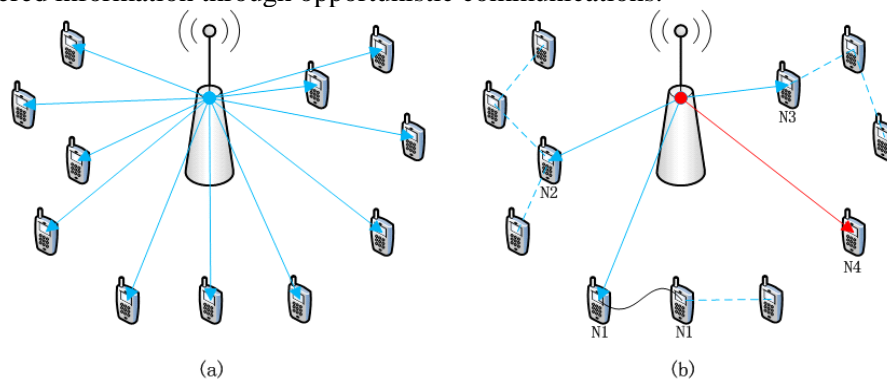


Fig. 1. (a) The infrastructure-only mode; (b) the opportunistic communication mode

The existing works selecting the target-nodes that merely have the highest probability to encounter other users as the initial sources, but they do not consider the message is fully delivered in the meeting time, also did not consider the messaging success is exceed the time can be tolerated. Meanwhile, the message dissemination through target-nodes on encounter spread will cause a huge amount of copies. In addition, they all utilize the historical trajectory information to calculate the sum of the encounter frequency between nodes, therefore they can

not avoid the situation that historical data couldn't reflect the current moving rule of human mobility.

We introduce NodeRank algorithm based on the regularity of human mobility ranks the importance of a node in the contact graph. The existing research results show that the mobility of human follows a certain spatial-temporal regulation and there is a potential predictability. So to take advantage the rule of human mobility can guide the opportunistic communications. NodeRank is inspired by the PageRank algorithm [2] used in Google's search engine to measure the relative importance of a Web page within a set of pages. Similar to the PageRank, NodeRank identifies the most popular nodes to disseminate the message, given that popular nodes are more likely to meet other nodes in the networks.

We study the target-nodes selection problem as the first step towards bootstrapping the cellular traffic offloading for information delivery in cellular networks. Our first contribution is that we proposed NodeRank algorithm to choose target-nodes and disseminate message. The NodeRank utilize the characteristics of human mobility which combined with the encounter frequency, contact time and inter-contact time. Our second contribution is to exploit the regularity of human mobility and apply the target-nodes identified by using mobility regularity to information delivery in the future. In the algorithm, we choose the target-nodes in consideration of the contact time and inter-contact time that to ensure the integrity and availability of message delivery. Our third contribution is using the NodeRank algorithm which combines the centralization and distribution. This algorithm chooses the target-nodes not only in consideration of the historical information of human mobility but also the present state of human mobility. With the historical data, we can ensure the practicability of prediction towards human mobility and with the present information, we can avoid the situation that historical data couldn't reflect the current moving rule of human mobility.

This paper uses the real-world data sets that are based on the encounter of Bluetooth to evaluate the effect of data traffic offloading. The trace-based evaluation shows our method offloads almost 75% data traffic compared with the traditional method without offloading data, although we also take the integrity and availability of message delivery into consideration by adding contact time and inter-contact time. Similarly, the offloading efficiency has improved up to 6% compared with Greedy algorithm, and the network redundant copies also have reduced almost 20%. The result of turns out that our method can well balance the success rate of message delivery and the network copy redundancies.

The rest of this paper is organized as following: Section 2 reviews related work on cellular traffic offloading. The NodeRank algorithm and target-nodes selection are presented in section 3. In Section 4, we evaluate the NodeRank algorithm performance through real-world mobility trace, and then conclude this paper.

2. Related Work

2.1 Cellular Traffic Offloading Technical Analysis

To alleviate the data overloading of cellular networks and improve the user experience, most mobile operators have introduced and started to implement mobile data offloading strategy [3]. To the best of our knowledge, we analyze and summarize the existing mobile data offloading solutions.

- Base station: The traditional approach of scaling network capacity with installing more base stations per area is always available, but not cost-effective and viable considering the pace at which the demand for data services is increasing. Moreover, from the perspective

- of energy saving, this will greatly improve the entire network energy consumption.
- Wi-Fi: Due to degradation of cellular services in overloaded areas, an increasing number of users are already using Wi-Fi to access Internet services for better experience. Wi-Fi is attractive because it allows data traffic to be shifted from expensive licensed bands to free unlicensed bands (2.4GHz and 5GHz). Nevertheless, the operator loses visibility (and hence control) of its subscribers whenever they are on the Wi-Fi network. Meanwhile, the operator is unable to deliver any subscribed content, leading to potential loss of revenue.
 - Femtocells: Femtocell is a small cellular base station typically designed for indoor use (e.g. in a home or office). It connects to the service provider's network via broadband (e.g. digital subscriber line, DSL), especially in areas where access would otherwise be limited or unavailable. Although femtocells can provide a highly effective method of easing the traffic carried by a macrocellular network, the careful planning is required as they operate in costly (licensed) and limited spectrum bands.
 - Spread spectrum: The development of spread spectrum technology improves the utilization of the frequency spectrum to a great extent and at the same time reduce the impact that comes from the co-channel interference, which greatly improves the capacity of the cellular network and can provide users with higher quality of the service. However, the spectrum resources are always limited and therefore the optimum utilization is restricted.
 - Device to Device: D2D communication technology [4] can improve the utilization of the spectrum resources and the network capacity, alleviate the burden of the network, reduce the consumption of the battery on the mobile terminals and increase the bit rate by obtaining the frequency resources and transmission power. However, when this technology sharing wireless resources with the cellular network, there will be some interference. Different from the opportunistic communications that D2D is needed during the communication under the control of the macro-cellular base station, means that need money cost.

There are also some methods such as Media optimization solutions, IP Flow Mobility and so on. Compared to the above methods, we mainly focus on offloading the cellular traffic from the perspective of system and network to optimize the quality of service, which is the research method, cellular data offloading via opportunistic communication, of this paper. Opportunistic communications is a promising solution to partially solve cellular traffic overloaded, because there is scarcely any monetary cost for it, and does not exist the shortcomings of the above methods. We will discuss the related work in the next section.

2.2 Cellular Traffic Offloading Through Opportunistic Communications

Using opportunistic communication to offload cellular traffic for mobile content dissemination applications is a novel and interesting idea and it has drawn great attention of the researchers. Opportunistic communication [5-6] is a technique that allows users without permanent connection to exchange information using low-cost proximity-based connection, such as Wi-Fi or Bluetooth, when they encounter opportunistically. Although using opportunistic communication to offload cellular traffic can significantly reduce the cost of communication, but it can only be applied to non-real-time messages that have high delay-tolerant characteristic and sometimes the messages will be unreachable for few nodes. Meanwhile, it will also generate more copy of the message.

The work [7] studies the trajectory of 100, 000 anonymized mobile phone users whose position are tracked for a six month period and the result shows that the mobility of human follows the characteristics of the periodic and return the location of high frequency to visit.

follows a certain kind of regularity. The work [8] indicates that there is a potential 93% average predictability in user mobility, an exceptionally high value rooted in the inherent regularity of human behavior. Consequently, to take advantage the rule of human mobility can guide the opportunistic communications. Nelson et al. proposed encounter-based routing [9] to encounter communication opportunity. Some studies [10-11] exploit unicast routing to enable a source to deliver the object to multiple destinations in an opportunistic network. All the above studies consider an environment where the source nodes are given. However, our work differs in that it focuses on how to choose a set of sources to disseminate the information to all subscribers before the information deadline.

The work [12] selects target-nodes to deliver the message by using the Greedy algorithm on the basis of the encounter probability of the users' historical trajectory and then disseminate the message to all the subscribed users. However, the selection of the target-nodes only considers the encounter frequency, at the same time the message dissemination through target-nodes on encounter spread will cause a huge amount of copies. Based on the study [12], the work [13] evaluates the amount of data exchanged from encountering to separating through Bluetooth between two phones. It shows that we need to consider the integrity of message delivery, however, the factor does not be applied in the algorithm. However, the factor does not be applied in the algorithm. The work [14] selected the target-nodes based community. Similar to the above mentioned works, it only uses the encounter frequency and the historical trajectory information to calculate the sum of the encounter frequency between nodes. On the basis of the node which has the maximum encounter frequency and by using the breadth first search algorithm, it will traverse the detection community. Then it chooses the target-nodes whose number equals to the number of the community and use the target-nodes to disseminate information.

Compared with the above studies, we add the contact time when we select the target-nodes, thereby guaranteeing the integrity of the data transmission. At the same time, we added to the contact time also better accord with the process of data transmission in opportunistic communication. In addition, we add the inter-contact time in NodeRank algorithm to ensure the availability of data delivery and received good results. In addition all the above works exploit the historical trajectory information to predict the present important nodes and the encounter of nodes for message dissemination. Using the global historical encounter information between nodes to establish contact graph for selecting target-nodes, although in reference [11] they observe that the historical data (one week or month) is enough to detect users that keep their importance, since the random of human mobility, for example a long distance travel or a travelling on official business, the node will disappear from the zone where it supposed to appear. So we adopt the combination of historical and present mobile information to choose the nodes and predict the encounter of the nodes, thereby disseminating the information. That is our NodeRank algorithm which combines the distributed and the centralized. We will discuss it in detail in next section.

The work [15] offloads the cellular traffic through opportunistic communication. It mainly proposes a Subscribe-and-Send architecture which consists of an application and an opportunistic routing protocol. Compared with spreading message by choosing target-nodes based on the analysis of the regularity of the human mobility, this architecture mainly emphasizes on the content distribution mechanism with base station instead of disseminating messages with the analysis of users' encounter regularity, which is different from using opportunistic communication to offload cellular traffic for mobile content dissemination applications.

3. Algorithm Description and Target Nodes Selection

Our focus here is on offloading cellular data traffic through opportunistic communications, especially in metropolitan areas with high population density. The content service providers who can tolerate large delay through the base station send the message to the target-nodes selected according to NodeRank algorithm and then the target-nodes can further disseminate the message to all subscribed users as large as possible.

3.1 NodeRank Algorithm

NodeRank algorithm is inspired by the most famous web page ranking employed by Google to rank web pages. PageRank measures the relative importance of a page within a web graph. Motivated by the success of this algorithm, we propose to apply a similar algorithm which we call NodeRank to rank the nodes in a contact graph. The main idea is that nodes with a higher NodeRank value will generally be more important nodes to disseminate the message, given that popular nodes are more likely to meet other nodes in the networks.

Before going into the details of our NodeRank algorithm, we revisit the properties of the PageRank algorithm. PageRank performs a random walk on the World Wide Web graph, where the nodes are pages and the edges are links between pages. PageRank gives the probability distribution used to represent the likelihood that a person randomly clicking on links will arrive at any particular page. PageRank is given by the following equation:

$$PR(p_i) = \frac{1-d}{n} + d \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)} \quad (1)$$

Where p_1, p_2, \dots, p_n are the pages, $M(p_i)$ is the set of pages that link to p_i , $L(p_j)$ is the number of outbound links on page p_j , n is the total number of pages, and d ($0 < d \leq 1$) is a damping factor which is defined as the probability, at any step, that the person will continue clicking on links instead of opening another random page. Various studies have tested different damping factors, but it is generally assumed that the damping factor will be set around 0.85 [2].

Firstly, we will introduce initial NodeRank algorithm namely iNodeRank. It uses the same idea as PageRank algorithm to represent the importance of a node. In the same way web pages are linked, we build a contact graph between nodes when they contact with each other. As we all know PageRank calculates page ranking value through directed graph which built by the citations between pages, nevertheless in most cases the important pages have connection with each other, so that is a bidirectional link between A and B. For the same reason, the encounter of nodes is equivalent this specific bidirectional link. Our iNodeRank algorithm was based on the characteristic of the bidirectional link to build the contact graph and to compute the importance of nodes. So we denote such a contact graph $G_c = (N_c, E_c)$ with a node set N_c and an edge set E_c . An edge $(N_i, N_j) \in E_c$ there is a contact between nodes N_i and N_j . Consequently, the iNodeRank value is given by the following equation:

$$iNR(N_i) = \frac{1-d}{n} + d \sum_{N_j \in E(N_i)} \frac{NR(N_j)}{L(N_j)} \quad (2)$$

Where N_1, N_2, \dots, N_n are the nodes (users), $E(N_i)$ is the set of nodes that encounter to N_i , $L(N_j)$ is the number of nodes encounter to N_j .

In iNodeRank algorithm we use only the encounter frequency between nodes to rank the important of nodes. Therefore, we choose the important nodes that have high ranking value to disseminate message through opportunistic communication with the aim of offloading traffic because they have more communication opportunities. But in the actual message transfer, except for encounter frequency having an impact on communication opportunity of message, the contact time and inter-contact time are also should be considered.

The contact time between two nodes will affect the integrity of the delivery of the message. That is to say, if a message needs 10 seconds to be fully delivered, however, the contact time between two nodes is only 5 seconds, thus the message transfer is not fully completed. Similarly, if the inter-contact time between two nodes is so long, for example over 12 hours or even longer, that exceed the maximum tolerate time then even the message is successfully delivered it is meaningless.

Therefore, we introduce NodeRank algorithm based on iNodeRank which takes contact time and inter-contact time as weight of the edge to ensure the integrity and availability of message delivery. The weighted NodeRank value is given by Eq(3).

$$NR(N_i) = \frac{1-d}{n} + d \sum_{N_j \in E(N_i)} \frac{NR(N_j)\omega(N_j, N_i)}{L(N_j)} \quad (3)$$

In Eq(3), $\omega(N_j, N_i)$ is the weight of the edge between node N_j and N_i , which integrates contact time (ω_{ct}) and inter-contact time (ω_{ict}) between the two nodes. $\omega(N_j, N_i)$ is calculated as follows:

$$\omega(N_j, N_i) = \omega_{ct}(N_j, N_i) * \omega_{ict}(N_j, N_i) \quad (4)$$

From Eq(2) and Eq(3) we can find that the three indexes encounter frequency, contact time and inter-contact time have close relations to the link between two nodes, in other words, the encounter of nodes. However, the three indexes are independent with each other in terms of computing relationships. With regard to encounter frequency, the more encounter with other nodes, the higher the ranking value will be. Combine with the above analysis on the integrity and availability of message delivery, we can know that the ranking value increased with the increase of contact time and decreased with the increase of inter-contact time.

The work [7] studies the opposite direction walking considering the moderate human walking speed is around 1 m/s and the contact duration of two mobile phones is about 20 seconds. Apart from the time used for detecting and connecting the devices, the average number of transferred bytes is 563.25 KB and the average duration of data transfer is 12.02 seconds which means within 29.8s the transferred bytes is 1MB has been higher than that of most of message. So we use 29.8s as the segmentation of contact time to indicate the integrity level of the data transmission. As for the inter-contact time, we define 12h as the interval. This is because if the inter-contact time is bigger than this value there will be excessive delay and the message will become meaningless.

The techniques we described thus far are suitable for centralized implementations where the contact graph is known a-priori. Clearly this may not be feasible in the mobile wireless environment we are considering. We exploit the combination method of distributed and centralized to select the target-nodes in the next section.

3.2 Target Nodes Selection for Message Dissemination

The combination of distributed and centralized of NodeRank is shown in Algorithm 1. In this

method, whenever two nodes meet each other, they exchange two pieces of information: (i) their NodeRank values included the value of the day and week; and (ii) the number of contact nodes they have. Then, the two nodes update their NodeRank values using the formula given in Eq.3. Implicitly, this algorithm exploits the mobility and contact behavior of the nodes since the NodeRank value is updated every time the nodes meet. Note that in this algorithm frequently seen nodes update their NodeRank more often. In fact, the rank of the node which meets with more other nodes will increase rapidly.

Algorithm 1 *Combination NodeRank (i)*

Require $L(n_j) \geq 0$

$NR(n_i)_{day} \leftarrow 0$ $NR(n_i)_{week} \leftarrow 0$

while n_i is in contact with $\{n_j\}$, $j = 1, 2, 3, \dots, n, j \neq i$

if $n_j \in E(n_i)$ *then*

$send(NR(n_i), L(n_i))_{day}$, $send(NR(n_i), L(n_i))_{week}$

$receive(NR(n_j), L(n_j))_{day}$, $send(NR(n_j), L(n_j))_{week}$

$update(NR(n_i))_{day}$, $update(NR(n_i))_{week} \leftarrow Eq(3)$

end if

while $\exists m \in buffer(n_i)$ *do*

if $n_j \in subscriber(m)$ *and* $NR(n_j) = \max\{NR(n_j)\}_{week\ and\ day}$

$Delivered(m, n_j)$

end if

end while

end while

Note that this combined NodeRank algorithm, where the day and week are respectively the ranking value from the past days and the past weeks continuing until the current time, is mainly in response to the disappearance of nodes in the region as a result of sudden long-distance travel. This is because the ranking value of a week is relatively stable and can predict the importance of nodes well, but it fails to predict the unexpected events of nodes. In the contrary, the ranking value of a day can reflect the importance of the nodes of the current time better, thus the combination will predict the importance of nodes more accurately and then disseminate the messages. Here $\max\{NR(n_j)\}_{week\ and\ day}$ combines the ranking value of

a day and a week, however, it takes the week ranking value as the first indicator and divides it into high, medium and low levels. When the week ranking values are at the same level, we choose the target-nodes to disseminate message based on the day ranking values.

Once the target-nodes have been chosen, the target-nodes will continue disseminate messages to subscribed users. There is one problem we could not ignore which is the strategy of message dissemination, because it must offer the best trade-off between cost (number of message replicas) and rate of successful message delivery. This is because end-to-end paths between two communicating nodes are rarely available in opportunistic networks [16]. In such situations, the nodes might still copy and deliver messages to nodes that are more likely to

meet the destination. There are also several existing works [17-18] for information dissemination in wireless networks. They all dedicated to reduce the number of cost while increasing the success rate, and were confirmed with good results. Here we adopt the High Rank sole One (HRO) method in which we choose the continuous dissemination nodes on the basis of NodeRank algorithm to ensure the success rate of message dissemination. By choosing the node which has the biggest NodeRank value among the meeting nodes to continuously disseminate message, we can reduce the number of message replicas effectively. Since in this situation the message delivery still uses the way of opportunistic communication, there will be a small quantity of message replicas. Meanwhile, because of property of without permanent connection of opportunistic communication, sometimes the message is unreachable for few nodes. At this time we still need to send messages through the base station. However, the result of our simulation turns out that our method can well balance the success rate of message delivery and the network copy redundancies.

4. Performance Evaluation

4.1 Mobility Traces

We now evaluate the performance of NodeRank based opportunistic offloading using the real-world mobility traces from the Reality Mining project [19]. The Reality Mining trace was collected using Nokia 6600 smartphones carried by 94 users from the MIT Media Laboratory and Sloan Business School, from 2004-07-26 to 2005-05-05. The information in this trace includes call logs, neighboring Bluetooth devices, associated cell-tower IDs, etc.

For the traffic offloading efficiency analyses in this paper, we used a subset of the mobility traces data which are Bluetooth data collected from October, November and December, 2004, since compared with other months these three months' data are complete enough to support the research of this paper. Moreover, Bluetooth data we used are different from cell tower data which think two mobile users are in contact of each other if their phones are associated with the same cell tower. The reason why we choose Bluetooth data is that there is a good evidence to prove that the two users have a close connect if a user scanning another via Bluetooth. What's more, when forwarding message, we use Bluetooth or Wi-Fi as a way of information dissemination which is close connect compared with under the cell tower. Furthermore, in the environment where the data covering areas are relatively small, if we consider that under the same cell tower is equivalent to encounter, then the contact graph established according to this rule consists of many strongly connected domains, which will enlarge the scope of message dissemination greatly and get the inaccurate success rate of transmission.

From the section 3 we have already known that we disseminate messages by choosing nodes according to the rank value which combines two time period and the NodeRank algorithm which considers the index of contact time and inter-contact time is weighted. Next we will compare the related performance of the algorithms, since iNodeRank and Greedy algorithm selects target nodes all only based on encounter frequency, so when we compare the NodeRank and Greedy algorithm, also add iNodeRank algorithm to compare. In **Table 1** we use the three algorithms in different time period to calculate the important nodes. From the Table we can see that the node sets are different from each other, where the day is lasted 6 hours starting from 6 December, 2004 12:00PM and the week is lasted 7 days beginning from 1 December, 2004.

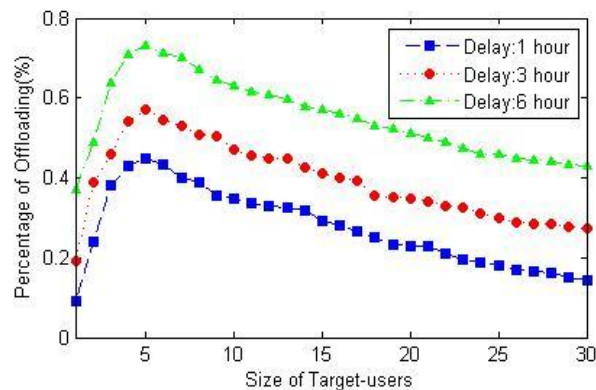
Table 1. The top 6 most active users for different algorithm

Algorithm	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6
Greedy(day)	30	91	7	102	46	90
iNodeRank(day)	46	84	7	83	30	24
NodeRank(day)	30	46	91	102	90	83
Greedy(week)	102	61	88	91	83	101
iNodeRank(week)	10	102	101	61	91	8
NodeRank(week)	102	10	61	101	91	88

4.2 Simulation Results

The performance of cellular traffic offloading is determined by two conflicting factors: (i) the success rate of message delivery; and (ii) the cost induced by the dissemination mechanism, i.e. the number of message replicas in the system. In the following, we assess the NodeRank performance with regard to these two performance indicators. In the meantime, we will compare the performance of Greedy, iNodeRank and NodeRank using the real-world traces in this section. Besides, in the experiment, we assume that all the users are subscribers.

Before analyzing the offload efficiency, we firstly consider the number of target-nodes selected to disseminate messages with the aim of getting better offload efficiency. As we known in above works have verified select the size of target-nodes is a NP-Completeness problem. In this paper, we utilize the real-world traces analyzes the problem of the size of target-nodes. We select the number of target-nodes is from 2 to 30, and respectively select 1 hour, 2 hours and 6 hours as message disseminate deadline. In Fig. 2 we can see the simulation results show that when using 5 target-nodes gained the best offloading effect in this data.

**Fig. 2.** Percentage of offloading for different size of target-nodes

Next, let us analyze the offload efficiency. We can see from Fig. 3 that the offload efficiency in unweighted iNodeRank algorithm is better than the weighted NodeRank. This is due to we taking contact time and inter-contact time into account in NodeRank algorithm that affected the delivery rate to some degree, but it ensures the integrity and availability of the message delivered. It is easy to understand that when using the iNodeRank algorithm, some of the messages failed to fully complete transmission, and some messages become meaningless once exceeding tolerant time.

However, the NodeRank algorithm which takes two indexes into consideration still can offload almost 75% data traffic in cellular network and the offloading efficiency has improved up to 6% than by Greedy algorithm. When the delay of message is less than 6h, it is obvious that iNodeRank and NodeRank are significantly better than the offload efficiency of Greedy

algorithm. Along with the growth of the delay time, the iNodeRank still maintains a high offload efficiency, NodeRank and Greedy converge to a consistent offload performance, which is not difficult to explain: when the increase of delay goes to a certain extent, nodes will encounter more and more users, even ones within the whole region, which corresponds with the phenomenon that the message spread to less users when it just begins to delay. Another question is about the value of the damping factor d , we are able to find out that the offload efficiency achieves the best value when d is around 0.85, in agreement with the PageRank algorithm [2].

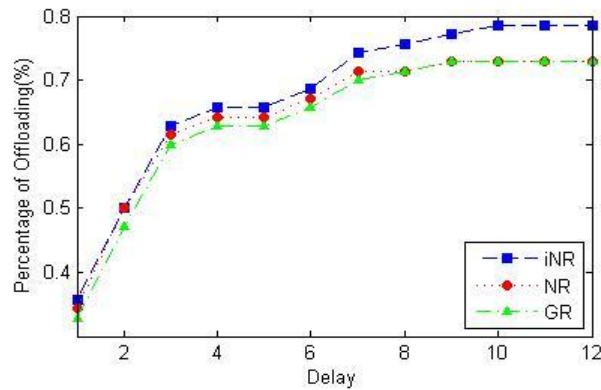


Fig. 3. Percentage of offloading comparison of iNodeRank, NodeRank and Greedy algorithms

We verified the offload efficiency of NodeRank algorithm and got the ideal result. However, it will cause the waste of a great deal of terminal resources and make the mobile phone users unwilling to cooperate if there are a large number of redundant copies. There are many papers which study how to establish effective incentive schemes to prompt people forwarding messages, because it is not the main research content of this paper, so we will not do too much discussion in this paper. From Fig. 4 we can see that since the Greedy algorithm disseminating messages through encounter dissemination, it has the maximum amount of copies. It is easy to see when comparing with iNodeRank, the number of copies generated by NodeRank is lesser. That is because, firstly, the dissemination method NodeRank and iNodeRank adopted which selects the nodes that have higher rank value to keep on disseminating messages when encountering a lot of nodes is different from Greedy algorithm. Secondly, the reason why NodeRank algorithm has few copies is that this algorithm is weighted, therefore the success rate decreases resulting in less number of copies.

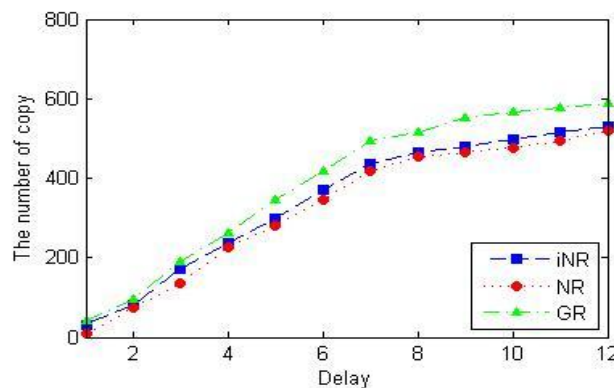


Fig. 4. The number of copy comparison of iNodeRank, NodeRank and Greedy algorithms

Then we extract 10 days a month randomly from October, November and December, every day we will start from 12:00 PM continuing 12 hours of delay time, through 10 randomly-selected experiments, the average offload efficiency obtained is shown as Fig. 5. We can see from the figure that iNodeRank still has the best offload efficiency, and NodeRank is slightly better than the Greedy algorithm, but because of the use of a longer delay time of 12 hours, so in November 2014 Greedy Algorithm has almost the same offload efficiency as NodeRank. Fig. 6 is that we use the same method as offload efficiency to calculate the copy quantity, we can find out from the figure the result still matches the previous copy simulation analysis.

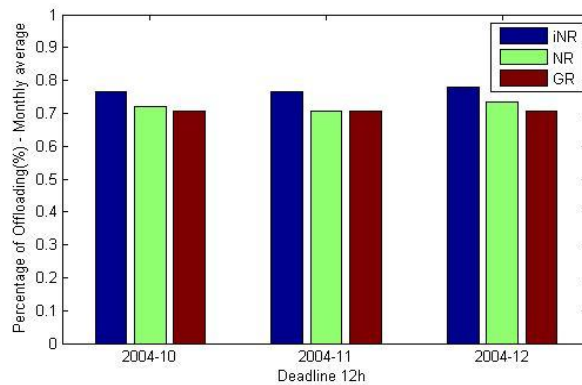


Fig. 5. Month average offloading efficiency comparison of iNodeRank, NodeRank and Greedy algorithms

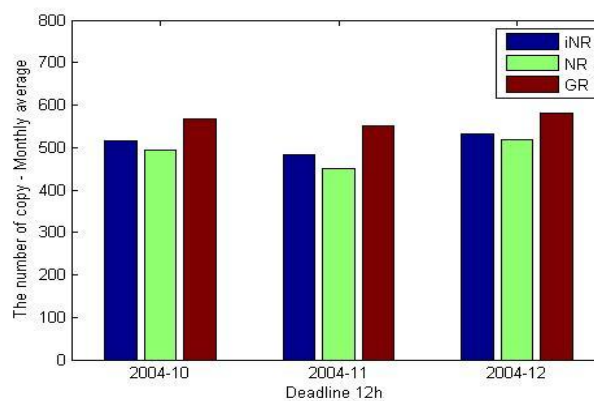


Fig. 6. Month average copies comparison of iNodeRank, NodeRank and Greedy algorithms

5. Conclusion

In this paper, we study the way to offload cellular network traffic through opportunistic communications. We propose NodeRank algorithm based on the regularity of human mobility to choose the target-nodes and disseminate information. NodeRank algorithm uses the encounter frequency characteristics to rank the important nodes, meanwhile we add contact time and inter-contact time to ensure the integrity and availability of message delivery to a certain extent. So from the perspective of the effect of data offloading, our improvement is limited. This paper uses the real-world mobility traces for simulation experiment to show that

compared with the traditional method without traffic offloading our method can offload almost 75% data traffic in cellular network. Similarly, the offloading efficiency has improved up to 6% than that by encounter frequency, and the network redundant copies also have reduced almost 20%.

References

- [1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update," 2014. [Article \(CrossRef Link\)](#)
- [2] S. Brin, L. Page, "The anatomy of a large-scale hypertextual web search engine," *Comput. Netw. ISDN Syst.*, vol. 30, pp. 107–117, 1998. [Article \(CrossRef Link\)](#)
- [3] A. Aijaz, H. Aghvami, M. Amani, "A survey on mobile data offloading: technical and business perspectives," *Wireless Communications of IEEE*, vol. 20, no.2, pp. 104-112, 2013. [Article \(CrossRef Link\)](#)
- [4] Kim S. Y, Lim C. H, Cho C. H, "Performance Analysis of a Dense Device to Device Network," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 9, pp. 2967-2981, 2014.
- [5] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott, "Pocket switched networks: Real-world mobility and its consequences for opportunistic forwarding," *University of Cambridge, Computer Lab, Tech. Rep*, 2005. [Article \(CrossRef Link\)](#)
- [6] M. Motani, V. Srinivasan, P. S. Nuggehalli, "Peoplenet: engineering a wireless virtual social network," in *Proc. of ACM MobiCom*, 2005. [Article \(CrossRef Link\)](#)
- [7] M. C. Gonzalez, C. A. Hidalgo, A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, pp. 779–782, 2008. [Article \(CrossRef Link\)](#)
- [8] Song C, Qu Z, Blumm N, et al. "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018-1021, 2010. [Article \(CrossRef Link\)](#)
- [9] S. Nelson, M. Bakht, R. Kravets, "Encounter-based routing in DTNs," *IEEE INFOCOM*, 2009. [Article \(CrossRef Link\)](#)
- [10] U. Lee, S. Y. Oh, K.-W. Lee, M. Gerla, "RelayCast: Scalable multicast routing in delay tolerant networks," in *Proc. of IEEE ICNP*, 2008. [Article \(CrossRef Link\)](#)
- [11] W. Gao, Q. Li, B. Zhao, G. Cao, "Multicasting in delay tolerant networks: A social network perspective," in *Proc. of ACM MobiHoc*, 2009. [Article \(CrossRef Link\)](#)
- [12] B. Han, P. Hui, M. Marathe, G. Pei, A. Srinivasan, A. Vullikanti, "Cellular traffic offloading through opportunistic communications: A case study," in *Proc. of ACM Chants*, 2010.
- [13] B. Han, P. Hui, V. Kumar, M. Marathe, J. Shao, A. Srinivasan, "Mobile Data Offloading through Opportunistic Communications and Social Participation," *Mobile Computing of IEEE Trans*, vol. 11, no. 5, pp. 821-834, 2012. [Article \(CrossRef Link\)](#)
- [14] Chuang Y J, Lin K C J. "Cellular traffic offloading through community-based opportunistic dissemination," in *Proc. of Wireless Communications and Networking Conference (WCNC)*, IEEE, pp. 3188-3193, 2012. [Article \(CrossRef Link\)](#)
- [15] LU Xiaofeng, HUI Pan, Pietro Lio, "Offloading Mobile Data from Cellular Networks Through Peer-to-Peer WiFi Communication: A Subscribe-and-Send Architecture," *China Communications*, vol. 10, no. 6, pp. 35-46, 2013. [Article \(CrossRef Link\)](#)
- [16] A. Mtibaa, M. May, M. Ammar, C. Diot, "PeopleRank: Social Opportunistic Forwarding," in *Proc. of INFOCOM*, 2010. [Article \(CrossRef Link\)](#)
- [17] C. Lindemann, O. P. Waldhorst, "Modeling Epidemic Information Dissemination on Mobile Devices with Finite Buffers," in *Proc. of SIGMETRICS 2005*, pp. 121–132, 2005.
- [18] M. Papadopouli, H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices," In *Proc. of MOBIHOC 2001*, pp. 117–127, 2001. [Article \(CrossRef Link\)](#)
- [19] N. Eagle, A. Pentland, D. Lazer, "Inferring Social Network Structure using Mobile Phone Data," in *Proc. of the National Academy of Sciences*, vol. 106, no. 36, pp. 15274-15278, 2009. [Article \(CrossRef Link\)](#)



Zhigang Li received the B.Sc. degree from Huanghe Science and Technology College in 2006 and the M.Sc. degrees from Beihang University in 2010. He is currently working toward the Ph.D. degree with State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications (BUPT), Beijing, China. His current research interests include mobile computing and cellular network.



Yan Shi received her Ph.D. degree from Beijing University of Posts and Telecommunications (BUPT) in 2007. She is currently a research staff of the State Key Laboratory of Networking and Switching Technology. Her current research interests include network architecture evolution, protocol design and performance optimization of future networks and mobile computing, especially mobility management technology.



Shanzhi Chen received his Ph.D. degree from Beijing University of Posts and Telecommunications (BUPT), China, in 1997. He joined Datang Telecom Technology & Industry Group in 1994, and has been serving as CTO since 2008. He was a member of the steering expert group on information technology of the 863 Program of China from 1999 to 2011. He is the vice director of State Key Laboratory of Wireless Mobile Communications, and the board member of Semiconductor Manufacturing International Corporation (SMIC). He has made great contributions to TD-SCDMA 3G industrialization and TD-LTE-advanced 4G standardization. He received State Science and Technology Progress Award of China in 2001 and 2012. His current research interests include wireless mobile communication, IoT and emergency communication.



Jingwen Zhao is currently pursuing her B.E in Beijing University of Posts and Telecommunications. Her major is e-Commerce Engineering with Law.