This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

Efficient Mobile Content Delivery Based on Co-route Prediction in Urban Transport

Le Shu Hongbo Jiang Xiaoqiang Ma Lanchao Liu Kai Peng Bo Liu Jie Cheng Yanbo Xu
Department of EIE, Wuhan National Laboratory for Optoelectronics
Huazhong University of Science and Technology, Wuhan 430074, China
Email: {slevenshu, hongbojiang2004, mxqhust, hustlanchao}@gmail.com,
{pkhust,liubo}@hust.edu.cn, {jiecheng2009,xuyanbohust}@gmail.com

***Abstract*—Routing is one of the most challenging open problems in pocket-switched-networks (PSN). In this paper, we propose a novel co-route media content forwarding scheme (CRMF), in which new contact opportunities are created for occasionally disconnected mobile users. Our study is inspired by two observations: one is that many people tend to make regular journeys to the same place, so their trajectories show a high degree of temporal and spatial regularity. The other is that the number of repeated journeys for an individual commuter is greater than that of the repeated contacts with another commuter who possess similar seasonal movement patterns. Our main contributions include: we properly install store-and-forward routers based on vehicle mobility patterns and human regular movement behaviors; we also propose a router-centric prediction scheme that collects passenger historical trajectory information to determine the delivery scheme. The simulation results demonstrate that this approach improves delivery ratio and also reduces the delivery latency compared to memory (history)-less delivery scheme.**

I. Introduction

People live in urban areas spend significant time on transport systems[6]. During these journeys to/from work, most commuters carry mobile devices with them which can communicate with other devices or with infrastructure within communication range through Bluetooth or Wi-Fi. Apparently, short-range device-to-device transfers have many advantages over infrastructure based solutions [3]. They do not require network coverage, and can be performed when the infrastructure is unavailable, or out of work. However, given that in such a device-to-device network, the end-to-end paths are intermittent and the topology is varying, the routing is a rather challenging issue. Considering the limited transmission rate between potable devices, the time to transfer large files might be much longer than human contact duration, thus it is needed to use cache-and-forward (CNF) routers with large storage [4], which provide opportunistic delivery for occasionally disconnected mobile users and in-network caching of content.

In this paper, we propose a CNF network infrastructure and a CRMF (co-route media content forwarding) scheme for multi-hop, peer-to-peer based content sharing on public transport. Our main idea is to use store-and-forward routers with large storage as data mules. The routers in our scenario are only equipped with short range wireless network interfaces. There are two kinds of routers, stationary routers and mobile routers (for routers installed in bus, we call them bus routers; for routers installed in station, we call them station routers). We use reliable hop-to-hop connection between nodes and routers as well as routers and routers to substitute for the intermittent connection between two nodes. When routers need to choose the next hop to forward the media content, our CRMF delivery scheme uses historical trajectory information: the copy of content will be delivered to the station router only if the destination node frequently passes by, and the station router will forward the content to the bus router only if the destination node frequently takes this bus route.

Our study is inspired by two observations from reality. These two observations are similar to which pointed out in[1] [2], however, with two different key insights. One is that each node has its own mobility schedule, and the dynamics of the inhabitants' daily mobility patterns show great spatial-temporal correlations[2]. The second observation is that journeys of an inhabitant probably follow the same route pattern day by day, however, the timetable may slightly change[1]. So the number of meeting with the 'familiar strangers [5]' for one commuter is less than his/her individual travel number. If we could create new contact opportunities for those nodes which are separated by time and space, the performance of forwarding content would be greatly improved. Due to the speciality of social interaction in public transport systems,we do not only relay on the direct contact of the passengers who possess similar seasonal movement patterns. Instead, we design a proper location-based infrastructure to provide indirect contacts between the nodes which move around the same place but not at the same time instance.

In this paper, we propose a network infrastructure for mobile content delivery based on location regularity patterns. Our main contributions include: first we properly install store-and-forward routers according to the vehicle mobility and human regular movement routes; then we explore the 'idle transit time' and propose an optimal delivery scheme for media contents between the routers.

The remainder of the paper is organized as follows, in section II, we introduce some related work. Section III portrays a public transport scenario and the context of our work. Section IV describes our approach, and explains the key aspects of the CRMF forwarding scheme in detail. Section V presents our simulation methods and results. Finally, the paper is concluded in section VI.

978-1-4244-5637-6/10/$26.00 ©2010 IEEE

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

II. Background and Related Work

Delay Tolerant Networks (DTN) have received significant research attention in the past few years [7], and most of the previous work focus on using opportunistic connections between nodes to spread information in a multi-hop way [8][9]. Knowledge about human connectivity and social interaction have been employed to improve the routing performance in multi-hop DTNs [10]. This section briefly introduces related work focusing on content sharing and routing in PSN networks.

In [11], storage only equipped with Bluetooth support, which is called bluespots, is placed on all buses in the public transportation system. Passengers can get their favorite media content from the bluespots, and the bluespots can renew their contents according to passengers' interest when the bus arrives at a bus stop. This system provides an approach to data distribution for mobile users in public transportation systems. BlueTorrent [12], proposed a way to effectively share content in spite of short link duration. Content is divided into a number of small pieces, and mobile users can exchange whatever pieces that are available. City commuters often stay together on the same bus or train for a large amount of time, so they could share content with fellow travelers through mobile device equipped with wireless technology [1]. Historical colocation and social information can be used to determine the best content source. This paper does not focus on the routing or remote delivery, but on the behavior of pair-wise interactions. Unlike previous work, we propose a network structure for media content in DTN networks and take advantage of vehicle mobility patterns and human regular movement behaviors to implement large-scale content exchange.

III. Scenario

A simple example shown in Fig.1 will serve to illustrate our proposed media content delivery scheme, and we make the following assumptions: people carry devices and move in the city as a part of normal schedule. They prefer to share particular content of interest with their friends as the journey may be long and boring.

Alice takes regular journey from station A to station B at her usual time. During her journey, she wants to deliver particular media files to her friends Bob and Carol. Bob is another commuter who travels to work along the same train line, but he is often half an hour later than Alice. Although they take the same bus/train route, there is no chance for them to meet each other and thus no direct communication between their devices; Carol is on her usual route from D to E at the same time as Alice, but they travel on different lines, so they cannot encounter each other during their journeys.

It will incur many drawbacks if we only rely on direct communication between two nodes. At first, the time to transfer media content can be much longer than human contact duration. Secondly, although people often follow seasonal movement patterns, these patterns may vary slightly. For example, Alice is required to be at her office at 8:30A.M., and she may get on bus between 7:45A.M. to 8:10A.M. in order to follow her job timetable. In this paper, we try to design a network infrastructure to ensure reliable hop-to-hop media content transfer for mobile users.

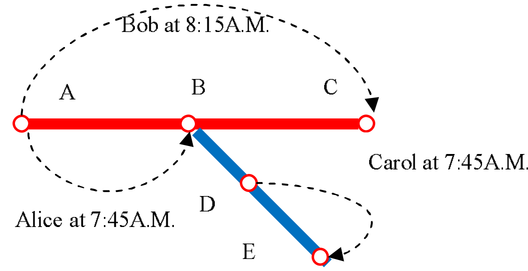

Fig. 1. An example of city commuter mobility model. The hollow circle denotes the bus/train station. The solid bold line stands for bus/train route. The dashed line denotes the mobility of city commuters.

IV. Approach

In this part, we first introduce the CNF network infrastructure to provide reliable hop-to-hop transport of media files for mobile users. We then describe how co-route patterns are recorded over time and used to estimate future co-route. After that, we illustrate how these predictions can be used by our relay process to maximize the chance of delivery rate.

A. CNF network model

The CNF network model is outlined in Fig. 2. Each stop and bus is installed with routers with large storage cache which can store media files in transit. We assume that all routers are equipped with short-range communication facilities, so our proposed network model maintains compatibility with mobile users and can be performed everywhere. And these two kinds of routers can only contact with each other when the bus parks at the station. During this period of time their contact time is long enough for media content sharing and can be exactly predicted.

Our routing algorithm is based on the following assumptions: the routers in bus/stop use wifi data transfers,which can occur at broadband speed [13], so the parking time is long enough for the media files transfer between bus routers and station routers. And the amount of time that travelers spend in their journey is long enough for them to exchange the media files with bus routers.

B. Co-route pattern

From the above analysis, we have designed a CMRF network infrastructure to enable media content delivery between mobile users. We next propose a routing scheme to limit copies of media content. The design goal is to achieve high routing performance while retaining a low cost. The routers only forward copies of content to the routers which have the higher delivery probability than the other routers. To perform this scheme, each router must know the associated forwarding metric for each destination node. A simple technique to measure this forwarding metric is to keep a profile for regular commuters. Each bus route logs all its previous passengers,

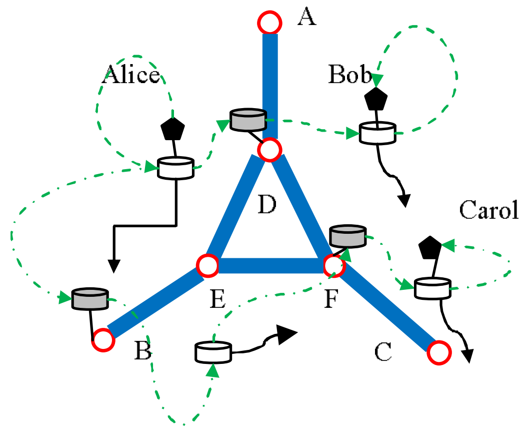

Fig. 2. Cache-and-forward network infrastructure. The solid line stands for the vehicle movement route, the dash line denotes the course of data delivery, and dark and light cylinders represent the static and mobile routers separately.

and records the mean frequency (the number of traveling) of every passenger. The higher the mean frequency, the greater the likelihood the passenger takes this bus route. However, this metric is less informative since the spatial regularity is ignored. In our approach, the bus router keeps a separate mean frequency of passengers for each station of a certain bus/train route. This detailed profile of the passengers mean frequency is kept in the form of bus profile, and all buses keep one copy of their bus profiles. Each station router also keeps a station profile for its regular passengers. And this station profile can be obtained directly from several bus profiles. Each station router first gets bus profiles from the buses passing by, then extracts all its regular passengers from those bus profiles, and finally records the mean frequency for each passenger and keeps detailed information in the form of station profile.

C. Algorithm

We now describe the steps of our co-route distributed forwarding protocol, in which historical co-route information is collected to determine the best forwarder. The following assumptions are made: commuters tend to make regular journeys over and over again, and their mobility patterns do not vary with time in a long term; routers in our network already hold a copy of their own regular traveler profiles.

When Alice enters the urban transport system and she wants to deliver a media content to Bob. As soon as she gets on the bus, she delivers one copy of content to the bus. When the bus runs along its route, it has to decide whether to forward a copy to stations during its following route. According to the forwarding metric mentioned before, if the mean frequency for Bob is high in the station profile, it can be inferred that Bob has a high probability to pass by the station in the near future, so the bus can forward the copy of content to the station router to maximize the delivery probability. If the content has not been delivered during the whole route, a copy of content will be forwarded to the last station router. A more formalized expression of the bus router forwarding scheme is shown in Algorithm 1.

Algorithm 1 Bus router forwarding scheme

Parameters: P denotes the sets of packets stored in the bus($p \in P$). $p.d$ contains the destination of packet p and $p.r$ is the remaining hop-count of packet p. Vector S ($S[d]$ = mean frequency for host d at station S) stores stop router profiles for regular passengers.

```
for all p ∈ P do
    if p has not been forwarded then
        if p.r > 0 and S[p.d] > THRESH then
            forward the packet p to the station S with
            ATTEMPT_RATE;
            p.r − −;
        end if
    end if
    if S is the last stop in the route then
        if p has not been forwarded and p.r > 0 then
            forward the packet p to the station S with
            ATTEMPT_RATE;
            p.r − −;
        end if
    end if
end for
```

When station router receives a copy of media content, it has to decide whether to forward the copy to the buses passing by. The station router will check whether Bob is a regular traveler of the bus router. If so, it will forward one copy to the bus. If the station is the last station in the whole route and the content has not been delivered to any bus, the station router will forward one copy of media content to all bus routers passing by. The detail of this station router forwarding scheme is shown in Algorithm 2.

Algorithm 2 Station router forwarding scheme

Parameters: P denotes the set of packets stored in the station($p \in P$). $p.d$ is the destination of packet p and $p.r$ is the remaining hop-count of packet p. Matrix S ($B[s][d]$ = mean frequency for host d on bus at station s) stores stop router profiles for regular passengers.

```
for all p ∈ P do
    if p has not been forwarded to any bus in a route then
        if p.r > 0 and ∑B[][p.d] > THRESH then
            the station forwards the packet p to the bus
            with ATTEMPT_RATE;
            p.r − −;
        end if
        if host p.d is on the bus then
            the station forwards the packet p to the bus
            with ATTEMPT_RATE;
            p.r − −;
        end if
    end if
end for
```

Our forwarding algorithms are distributed and they exploit resource conservative policy. Each router only needs to keep record of local regular passengers, rather than the whole routing information and seasonal patterns of all passengers. Mobile hosts do not have to keep any record of the social interactions, which helps to reduce the storage consumption and computational burden.

V. Evaluation

In this section, we evaluate our CMRF approach through several experiments. We begin with an description of our dataset. Then we describe the simulation methods and settings, and finally we discuss the results of the simulations in detail.

A. Dataset

The transport topology we use in our experiment is as shown in Fig. 2. There are six bus stations and six bus routes (A-D-E-B, B-E-D-A, B-E-F-C, C-F-E-B, A-D-F-C, and C-F-D-A). We assume that all buses are scheduled to depart from station A, B or C every 45 minutes to one hour, and pass by each station every 10 or 15 minutes. In other words, there are 18 buses running on a single route each day from about 6A.M. to 23P.M.. We randomly generate the source and destination station for each regular traveler, and we generate the regular traveling time for commuters with fluctuations in a certain range. Each commuter will either travel to/from work along his usual route , or take the day off from work. The whole dataset obeys the following property which is pointed out in [2]. For each day, there are obvious peak hours of bus trips, and the volume of traffic in these hours occupies almost half that of the whole day. This situation reflects the daily life pattern that most commuters begin their work around 9A.M. and go back home around 18P.M..

B. Simulation methods and settings

We will compare our method with other two ones of content delivery in order to validate the effectiveness of ours. **1. Random** - routers deliver copies of media files to their current neighbors at random. **2. Flood** - routers forward copies of media files to any possible adjacent routers and guarantee a maximized delivery rate. All the methods are based on the hop-count-limited forwarding [10]. And these protocols differ only on the delivery step, with all other steps being the same.

The default setting in our simulation is show in Table 1. The absent rate means the probability that a commuter takes a day off from his/her regular route. In each simulations, we vary one of the four variable parameters as shown in the table. At the beginning of the simulation, every node sends 20 media files to 20 randomly selected destination nodes. The total simulation time in each experiment is 5 days (120 hours). A message is considered to be successfully delivered if one of its copies has been sent to its destination before the end of the simulation.

C. Result

This section presents the simulation results of the three methods, and discusses how the simulation parameters affect the network performance. We first study the temporal distribution of delay and the result is shown in Fig.3. We randomly selected 2000 commuters as the default setting and measured the time consumption for them to transfer the media files. We can see that in our proposed scheme, 66% commuters can successfully deliver media files to their friends in 20 hours and less than 10% commuters need more than 40 hours. As a comparison, it takes much more time to transfer media files in the other two schemes.

TABLE I
SIMULATION PARAMETERS

Parameter name	default	range
Number of commuters	300	$100 \sim 600$
Attendance rate	0.6	$0.1 \sim 0.8$
Initial hop-count(H)	2	$1 \sim 6$
Absent rate	0.1	

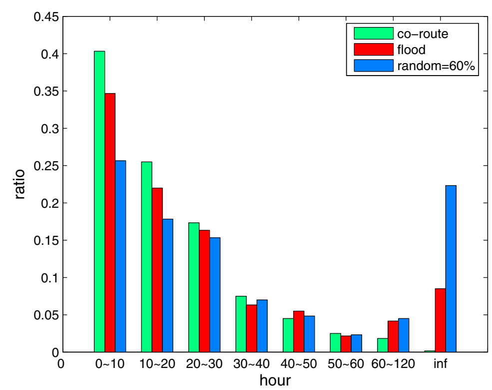

Fig. 3. Delay distribution

Then we study the affect on our performance metrics (which include delivery rate, delay and buffer cost) with varying initial hop-count and attendance rate. Fig. 4(a)-(c) shows that for all three delivery methods, the success rate increases as the initial hop-count. Our scheme can asymptotically approach 100% with the initial hop-count 3, and the number of total copies in station and bus routers is 634, 12% better than Flood and 27% better than the Random method. The delivery rate in Flood needs 6 initial hop-count to achieve 100%, and the total cost is 1364 copies. Our average delay is also the least among the three methods.

The change of delivery rate, delay and the cost with different attendance rate is shown in Fig. 4(d)-(f). The higher the attendance rate is, the more likely the commuters possess similar seasonal movement patterns. When the attendance rate is low, which means inhabitants seldom use public transport or rarely keep to their job timetable, our method has no advantage over the other two methods. However, as the attendance rate increase to 30%, our co-route approach is 7% better than Flood and 20% than Random in delivery rate. The average delay gets lower as the attendance rate increase. When the attendance rate

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

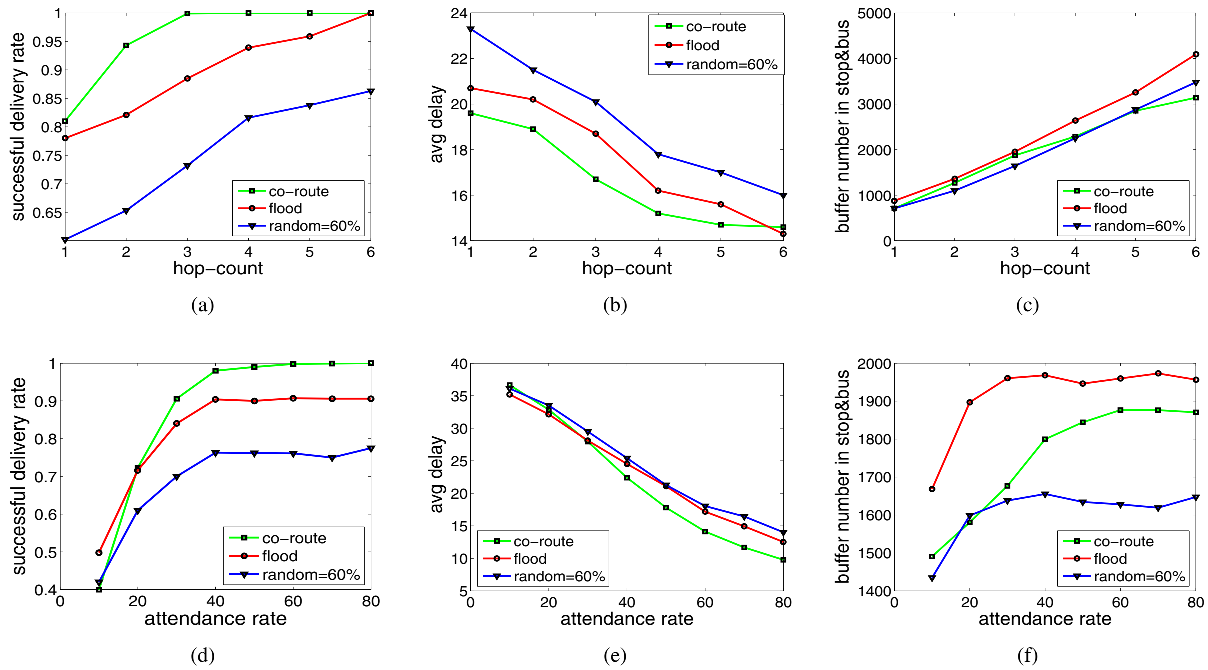

Fig. 4. Delivery rate, delay and buffer cost verse number of hop count and attendance rate

achieves 80%, the delay is only 9.78 hours for our proposed scheme, which is 2.8 hours less than Flood and 4.3 hours less than Random.

We also study the affect caused by increasing the number of commuters. We increased the number of commuters from 100 to 600, and observed that there is no significant change in delivery rate as well as average delay, and the number of copies increases roughly linearly. Our scheme shows better stability over the other two methods. Due to page limits, the result is not shown.

All the above simulation results confirm that compared with two existing algorithms, our algorithm performs better in delivery rate, delay and buffer cost.

VI. Conclusion

In this paper, we propose a novel CMRF network structure for media content delivery in urban transport. Our approach takes advantage of vehicle mobility patterns and human regular movement behaviors to install large storage routers, which act as data mules to transfer content from one place to another. We also design a memory-based forwarding scheme to ensure the routing performance. Our future work may focus on evaluating our proposed CMRF on real world datasets.

Acknowledgment

This work was supported in part through Chinese National 863 project (No.2007AA01Z223), National Natural Science Foundation of China (No.60803115, No. 60873127), Chinese National University Basic Research Funding (No.M2009022), the Youth Chenguang Project of Wuhan City (No. 200850731351, No.201050231080), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry. The corresponding author is Hongbo Jiang.

References

[1] L. McNamara, C. Mascolo, and L. Capra, "Media sharing based on colocation prediction in urban transport", In *Proceedings of MobiCom,* 2008.

[2] L. Liu, A.Biderman, and C.Ratti , "Urban mobility landscape: real time monitoring of urban mobility patterns", In *Proceedings of Computers in Urban Planning and Urban Management (CUPUM),* 2009.

[3] J.Scott, P. Hui, J.Crowcroft, and C.Diot, "Haggle: a networking architecture designed around mobile users ", In *Proceedings of Annual Conference on Wireless On-demand Network Systems and Services,* 2006.

[4] S. Paul, R. Yates, D. Raychaudhuri, and J. Kurose, "The Cache and forward network architecture for efficient mobile content delivery services in the future Internet", *Innovations in NGN: Future Network and Services, First ITU-T Kaleidoscope Academic Conference,* 2008.

[5] E. Paulos and E. Goodman, "The familiar stranger: anxiety, Comfort, and play in public places", In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI),* 2004.

[6] G.Lyons and K.Chatterjee, "A human perspective on the daily commute: costs, benefits and trade-offs", *Transport Reviews,* Vol.28, pp.181-198, 2008.

[7] K. Fall, "A delay-tolerant network architecture for challenged internets", In *Proceedings of ACM SIGCOMM,* 2003.

[8] A. Vahdate and D. Becker, "Epidemic Routing for partially connected ad hoc networks", *Technical Report, Duke University,* 2002.

[9] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Spray and wait: an effcient routing scheme for intermittently connected mobile networks", In *Proceedings of ACM WDTN,* 2005.

[10] C. Liu, and J. Wu, "An optimal probabilistic forwarding protocol in delay tolerant networks", In *Proceedings of ACM MobiHoc,* 2009.

[11] J. LeBrun and C. Chuah, "Bluetooth content distribution stations on public transit", In *Proceedings of Workshop on Decentralized Resource Sharing in Mobile Computing and Networking,* 2006.

[12] S. Jung, U. Lee, A. Chang, D. Cho, and M. Gerla, "BlueTorrent: cooperative content sharing for bluetooth users", In *Proceedings of IEEE PerCom,* 2007.

[13] J. Eriksson, H. Balakrishnan, and S. Madden, "Cabernet: Vehicular Content Delivery Using WiFi", In *Proceedings of MobiCom,* 2008.