

Exploring Passenger Dynamics and Connectivities in Beijing Underground via Bluetooth Networks

Rina Wu[#], Yuan Cao[#], Chi Harold Liu[†], Pan Hui[§], Li Li[†] and Erwu Liu[‡]

[#]Beijing University of Posts and Telecommunications, Beijing, China, [#]Lehigh University, PA, USA

[†]IBM Research - China, Beijing, China, [‡]Tongji University, Shanghai, China

[§]Deutsche Telekom Laboratories/TU-Berlin, Berlin, Germany

E-mail: [#]rina.l.wu@gmail.com, [#]yuc211@lehigh.edu, [†]{chiliu,lilicrl}@cn.ibm.com

[§]pan.hui@telekom.de, [‡]erwu.liu@ieee.org

Abstract—Unlike traditional opportunistic networks characterized by relatively very short connectivity period among mobile devices during their contacts, in this paper, we explicitly consider the underground/bus scenario where the inherited *structured* mobility pattern enforces more stable passenger behaviors. If with the aid of potential formation of Bluetooth-enabled multi-hop wireless networks, one can expect a type of new emerging applications for passengers on the move. In particular, we conducted a thorough experimental study on passenger dynamics and connectivities in Beijing Underground during peak hours, via building an information collection application in an Android-based smartphone. From the total of 41,806 records for 120 days, we have performed extensive analysis on a variety of statistics for the number of neighbors, the lifetime of the neighbors, the received signal strength indicator (RSSI) and its derived physical distance, flow speed of the nearby passengers, battery usage rate, etc. Then, we further prove our justification that Bluetooth is the perfect technology to build up a relatively small multi-hop wireless network with the network size of four nodes, and in particular it is applicable and may fundamentally drive new applications in an underground environment during the peak hour, where two notable examples are music album sharing and neighborhood gaming.

I. INTRODUCTION

The past several years have seen the astounding proliferation of affordable, wireless, and easily programmable mobile computing and communication devices such as smartphones and now, tablet computers. While integrated media and location-tracking features (e.g., cameras, GPS receivers, accelerometers, etc.) have become standard fare, one can expect a rapid increase of other “sentient” functions to vividly interact with nearby *neighbors* of the similar social interests/preferences. Notable examples include the “participatory sensing” [1], [2] (in which participants use devices to collect, possibly analyze, and make available nearby environmental data for large-scale applications), mobile file sharing [3], Micro-Blogging system [4], Bluetooth-enabled stranger finders [5], and Bluetooth-enabled poster sites in London Underground [6], etc.; however all above not fully exploit the features in underground/bus environments.

On the (almost) standard fare of the communication interfaces, 3G, Wi-Fi, and Bluetooth are three major enabling technologies. The resulted network is often referred as an *opportunistic* network, where contacts among users are usually opportunistic and connections among devices are intermittent. While on the one hand people argue that these uncertainties



Fig. 1. A random phone camera shot of the passengers in Beijing Underground during the peak hour, where one can see how crowded the train is and passengers are mostly playing around their handsets. Handsets are red-circled which can potentially form a multi-hop ad hoc network.

introduce significant challenges for designing any system running over it, it indeed offers new opportunities to exploit these dynamics and uncertainties as new information can be exchanged and shared. From the networking perspective, though, the problem of 3G is over used and prohibitively expensive subject to wireless data plans with finite usage limits and steep overage charges. The alternative is through a Wi-Fi network but that requires the deployment of access points nearby that needs credentials and cumbersome authentication processes.

Bluetooth [7], nevertheless, is partly distributed and partly centralized where devices are organized in groups denoted as “piconets”. In each group, a master node controls the transmissions to maximum seven other devices. Features include that all active nodes in a piconet share the same channel, communication to and from a slave device is always performed through its master, and a Bluetooth device can timeshare among different piconets. In particular, a device can be the master of one piconet and a slave in other piconets, or it can be a slave in multiple piconets. Devices with multiple roles will act as gateways to adjacent piconets, resulting in a multi-hop ad hoc network called a “scatternet”. Nevertheless, [8] proves

TABLE I
THE DATA FORMAT COLLECTED FROM THE ANDROID-BASED APPLICATION

Field 1	Field 2	Field 3	Field 4	Field 5
timestamp	MAC addr.	RSSI	device name	rem. battery (%)

that this organization is only applicable for relatively *small* ad hoc networks considering the complexity in establishing a connected topology.

In this paper, we aim to make full use of the Bluetooth network in an underground [6], [9] (see Fig. 1) or bus environment, where passengers follow a relatively “structured” mobility pattern given the fixed route of the transportation operations; and therefore, it is somehow guaranteed that the inherited connectivity period is at least one stop (e.g., two minutes on average in Beijing Underground). Interestingly, building on top of this, and *preference*-based social relationship, it is natural that we are attempting to bridge the Bluetooth-based multi-hop ad hoc network with its social circles. We therefore boldly foresee a set of potentially new applications to emerge, like music album sharing and neighborhood gaming, by exploiting the passenger dynamics and the resulted connectivities.

To this end, this paper presents a recent experimental study on passenger connectivities in Beijing Underground during peak hours, via building information collection application in an Android-based smartphone. From the total of 41,806 records, we did extensive analysis on a variety of statistics for the number of neighbors, the lifetime of the neighbors, the received signal strength indicator (RSSI) and its derived physical distance, flow speed of the nearby passengers, battery usage rate, etc.; and we have proved our justification that Bluetooth is the perfect technology to build up a relatively small multi-hop wireless network with the network size of four nodes, and in particular it is applicable in an underground environment during the peak hour.

The rest of the paper is organized as follows. In Section II, we establish a formal architecture of our system composed of the Android application and the companion back-end analysis system. Then, Section III describes the five performance evaluation metrics to be used in the experimental analysis in Section IV. Lastly, after presenting a few interesting emerging new applications in Section V building on top of the proposed Bluetooth network in the underground, in Section VI, we highlight related research activities. Finally, Section VII concludes the paper and presents the future work.

II. SYSTEM ARCHITECTURE

This section presents a formal system architecture to describe our data collection and analysis framework. We collect passenger data in the underground by our Android application built in our smartphone, as shown in Fig. 2, and the information, including the timestamp, MAC address, RSSI, device name, and remaining battery are collected and further recorded in a .txt file by this application, as shown in Table I.

We first built an Android application to search for the Bluetooth devices around the test phone, and Fig. 2

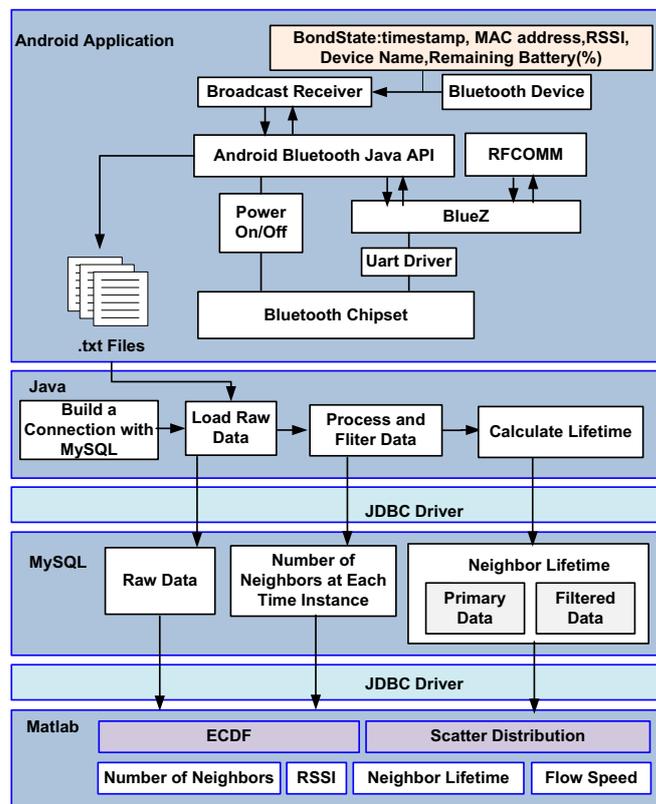


Fig. 2. Software architecture for our developed Android application, data loader and analyzer.

shows the architecture of this application. The application first detects whether there is a Bluetooth adapter inside the handset, and then check the Bluetooth power status through an Android Java API. Discovery process registers an `BroadcastReceiver` to receive the incident sent by function `sendBroadcast()`, which contains information of RSSI and a `BluetoothDevice` class. We extracted the device name, unique MAC address, and bond state from the `BluetoothDevice` class, and then combined them into a record file. Notice that, the next scanning process will not begin until the previous process finishes.

As for the data analysis part, we write a set of module to input the .txt files into the memory by Java language, and also further prepare the two data sets, primary and filtered for lifetime analysis. Then, JDBC driver is used to help load these data to the MySQL database, and create four tables, raw data table, lifetime table for primary and filtered data, and number of neighbor table. To further efficiently analyze the characteristics of our data mathematically, we leverage Matlab and use JDBC driver again to link it with the MySQL database. Then, to explore the distribution of the passenger behaviors, we write three separate analysis programs: one for the Empirical Cumulative Distribution Function (ECDF), and one for displaying the results in the scatter distribution function, and the last one to explore the flow speed of passengers, and we output those results.

III. PERFORMANCE METRICS

This section presents five performance metrics that we use to evaluate the statistical behavior of the passengers in Beijing Underground, namely: (a) number of neighbors nearby, (b) the lifetime of neighbors, (c) the RSSI and derived distance between the tester and the neighbors, (d) the passenger flow speed, and (e) energy consumption rate.

1) *Number of Neighbors*: We developed a function in Java to compute the number of neighbors, denoted as n , based on the collected data records at each time instance, as an important indicator representing the Bluetooth network size, and from which one is able to justify if Bluetooth, as a short-range and fast wireless communication technology, can satisfy our network formation requirements. Besides, the variability, or variance, of the number of neighbors at each time instance can show the dynamics of the targeted underground environment; e.g., if it is much deviated as time goes, then one can justify and possibly make a conclusion that the passenger mobility is relatively high in this scenario, but this is even more clear with respect to (w.r.t.) the flow speed of the number of neighbors discussed later.

2) *Neighbor Lifetime*: The lifetime of each neighbor, denoted as l , is computed as the time difference between the first time a particular MAC address (or the corresponding handset owner) appears and the last time it appears within one round of data sensing (i.e., before the tester completes his/her journey). We compute the lifetime for all MAC addresses appeared in our data records by this rule, but interestingly in the raw data set, we found that there are MAC addresses that only show once; and thus, for those records, we opt to adopt two different processing methods: one is to assume their worst case travel in Beijing Underground that a particular passenger's journey only lasts for one stop with the tester and then leaves the train, or, 120 seconds; while the other method is simply to filter out these data to give the best estimate. We believe that the lifetime metric can directly demonstrate how long a passenger can stay in the neighborhood of the tester, as a significant parameter for the network stability.

3) *RSSI and Distance*: The RSSI, or P_{Rx} , is sensed by our Android application representing the received signal power in dBm; and it implies the physical distance between the transmitter Tx and the receiver Rx within the proximity. Given the standard specification of Bluetooth physical layer parameters [10], we are able to calculate the free space path loss γ by:

$$\gamma = P_{Tx} - P_{Rx} = P_{Tx} - \text{RSSI}, \quad (1)$$

where P_{Tx} and P_{Rx} denote the transmission power and the receiver power in dBm, respectively. According to [10], we use $P_{Tx} = 2.5\text{mW}$ in Bluetooth Class2 radio standard, and then we have:

$$P_{Tx} - \text{RSSI} = 20 \log d + 20 \log f + 32.45, \quad (2)$$

where d denotes the distance between the Tx and Rx in kilometers and f is the signal frequency in MHz. Referenced from the official specification again, Bluetooth occupies the ISM

band from 2400 – 2800MHz [10], and we take $f = 2400\text{MHz}$ in our computation. We have:

$$d = 10^{\frac{1}{20}(P_{Tx} - \text{RSSI} - 20 \log f - 32.45)}, \quad (3)$$

or: $d = 10^{(-\frac{1}{20}\text{RSSI} - 3.3025)}$ m, where RSSI is input in dBm.

4) *Passenger Flow Speed*: It directly indicates the variability of the neighboring environment, as passengers come and go unpredictably. To compute the flow speed ν of the neighbors, we first figure out the same MAC address records in the two adjacent scanning time instances, and then calculate the number of the rest of MAC addresses, as the size of the change of dynamic passenger flow over time.

5) *Energy Consumption Rate*: It is undoubtedly arguable that the main functionality of a mobile phone is to make phone call, and/or send text messages; and thus we are interested in knowing how fast the battery is drained while performing the Bluetooth scanning. The parameter used for evaluation is the percentage of battery loss per second, denoted as r .

IV. EXPERIMENTAL RESULTS

In this section we describe our extensive experimental results in detail, based on our collected total 41,806 records for four consecutive months. As shown in Fig. 3(a), we first plot the ECDF w.r.t. the number of neighbors and compute its median value $\theta(n) = 4$, mean $\mu(n) = 4.6653$, and standard deviation $\sigma(n) = 3.2935$, representing that on average one can expect four handset owners already switching ON his/her Bluetooth. It is worth noting that this number is actually the worst case result as many people intentionally switch OFF their Bluetooth to save energy, as not many applications currently making full use of it. Therefore, we can safely conclude that the size of potentially formed Bluetooth ad-hoc network is relatively small, in accordance with what [8] has proved that as a short-range wireless communication technology, the complexity of forming a master-slave-based network can increase exponentially with the size of the nodes. Therefore, this experiment proves that our assumed scenario, underground/buses, are viable targets. Meanwhile, from the relatively high standard deviation $\sigma(n) = 3.2935$ compared with the mean, one can justify that high dynamics of the underground environment is as expected especially during the peak hour.

To further justify our findings of the dynamic environment in Beijing Underground, we analyze the lifetime of each neighbor as illustrated in Fig. 3(b). The characteristics of the lifetime distribution are displayed in its ECDF. The red one presents the results from the primary data set, i.e., assuming the 120s (or one stop) stay of those handset owners who appears only once, while the blue curve shows the result of filtered data set excluding the one-time appearance owners. We found that the mean value for filtered data set is $\mu_{\text{filtered}}(l) = 470.91\text{s}$, i.e., on average 3.924 stops, higher than the worst case of the primary data that $\mu_{\text{primary}}(l) = 279.46\text{s}$, or 2.329 stops. Furthermore, it is interesting to see that 50% of our filtered lifetime data are over 225s, i.e., more than half of the passengers one can meet every day will travel approximately two stops together, as seen in Fig. 3(b).

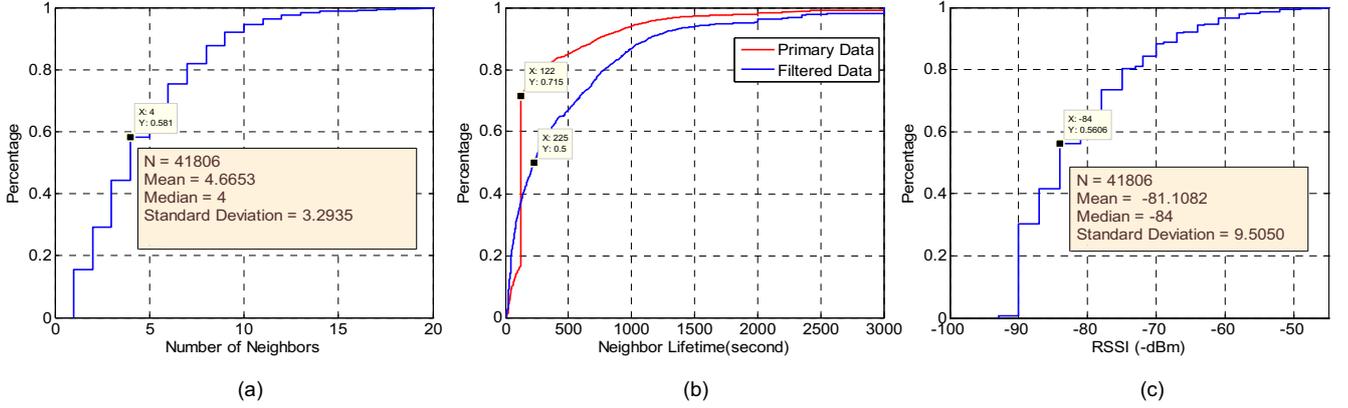


Fig. 3. Analysis results for the ECDF of (a) the number of neighbors, (b) the lifetime of neighbor, and (c) RSSI.

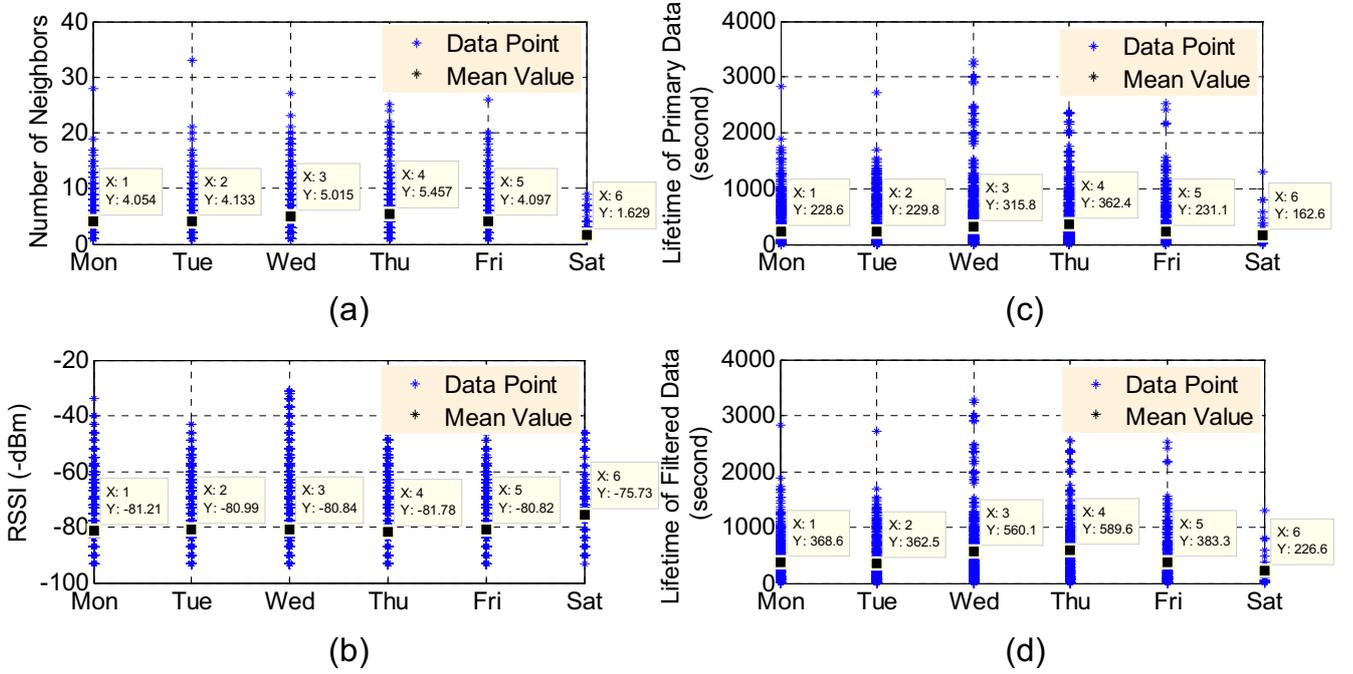


Fig. 4. Scatter distribution of (a) the number of neighbor, (b) RSSI, (c) neighbor lifetime for the primary data, and (d) neighbor lifetime for the filtered data, over a week.

Fig. 3(c) shows the ECDF w.r.t. the RSSI sensed by our Android application, where it can be seen that the mean RSSI $\mu(P_{Tx}) = -81.1082\text{dBm}$, and variance $\sigma(P_{Tx}) = 9.5050\text{dBm}$; and furthermore, given that $P_{Tx} = 4\text{dBm}$, we are able to compute the average Bluetooth radio transmission distance $\mu(d) = 5.6\text{m}$. According to the Beijing tube design principle, the length of the cabin is 20-22.5m, and thus half of the passengers in a cabin on average are within the communication range of the test phone.

Fig. 4 shows the scatter distribution of the number of neighbor, RSSI, and neighbor lifetime over a week. It is observed that except for the Saturday, the mean value for the number of neighbors on each working day is close enough to its overall mean $\mu(n) = 4.6653$, meaning that passengers go to work everyday by underground at the relatively similar pattern. The same result can be seen in Fig. 4(b), that the mean value of distributed RSSI everyday is around $\mu(P_{Tx}) = -81.1082\text{dBm}$, consistent with the ECDF shown in Fig. 3(c). Nevertheless,

their lifetime deviates dramatically from the mean and the range can even goes beyond 3000 seconds (or 50 mins), indicating that the network is indeed highly dynamic where a small probability does exist that some passengers follow the same way with the tester during the journey. Saturday sees lower passenger lifetime on average, possibly because the underground transportation is more used for leisure purposes on weekends and thus the destinations of the passengers are not as concentrated as working days (an example is that the City of Beijing, and Zhongguancun Science Park are the two of the hot spots aggregating much traffic every working day). All these results, again, have successfully justified our previous findings on the passengers dynamics in the Beijing Underground, but it shows its internal structure somehow as the two separate trends for working days and weekends.

Fig. 5 shows the distance distribution in ECDF, computed from the RSSI in (3). Fig. 6 shows its physical illustrations, where the color bar represents the distribution percentage

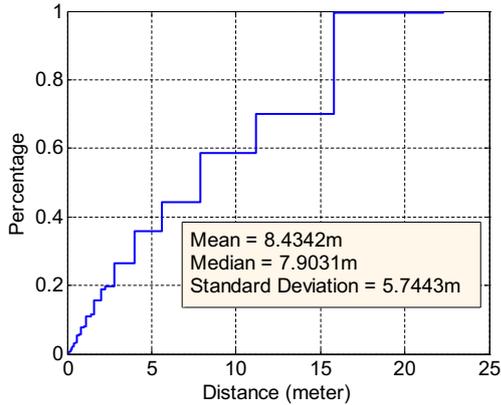


Fig. 5. The ECDF for the distance distribution.

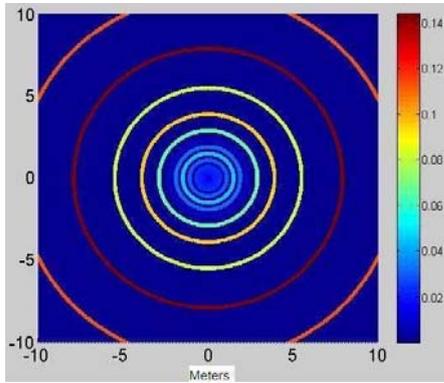


Fig. 6. An illustrative distance distribution from the tester and beyond, within 10 meters.

of the passengers, relative to the center position of our test phone. In our experiment, we found that the maximum calculated distance is $d_{\max} = 23\text{m}$, w.r.t. a very low RSSI that corresponds to a low Bluetooth datarate (and thus may not be applicable for neighborhood-oriented applications). This is in line with the above-mentioned Beijing Underground cabin design principle that the maximum cabin length is 22.5m, and thus Bluetooth signal strength deteriorates significantly from one cabin to another, or passengers in another cabin is not sensible. It is observed that the most dense area is within the circle of $d = 8\text{m}$ away from our test phone, and its ECDF up to 8m reaches 60%. Nevertheless, it is necessary to point out that this diagram just shows the worst case scenario, where passengers are not fully aware of the potential interesting Bluetooth applications within neighborhood. Also, our observation is not continuous due to the low accuracy of the sensed RSSI (as some corrupted bits in the signal).

Lastly, Fig. 7 shows the ECDF of the passenger flow speed indicating how dynamics they come and go over time, where we observe that on average one can expect $\mu(\nu) = 6$ people flowing in and out per minute, i.e., for one stop two minutes, one can expect 12 different new faces around. This result implies that potentially a massive amount of information can appear around the tester, and even more meaningful if considering their belonged social circles. Meanwhile, passengers are not allowed to move out of the train even though given the high dynamics, and thus the flow speed during one stop is unsurprisingly zero. Following this “zero to a-large-number”

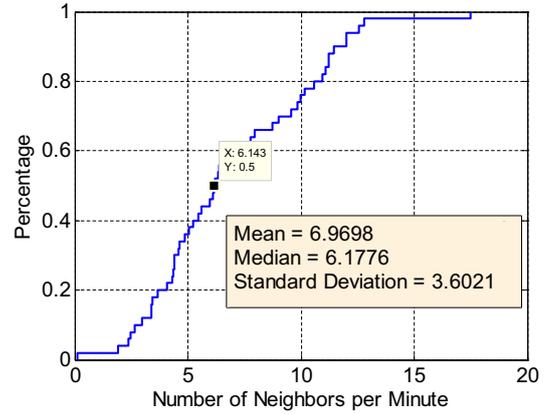


Fig. 7. The ECDF of passenger flow speed.

pattern, the inherited structured mobility pattern can be further derived, but left for the future work.

Finally, to complete our analysis, we also examine the battery usage rate while constantly scanning the Bluetooth neighbors. Surprisingly, we found only $\mu(r) = 0.00078576\%$ per second energy consumption rate on average, i.e., the energy consumption in discovery process can be negligible.

V. POTENTIAL APPLICATIONS

There are potentially three main streams of applications that will be benefited from these short range communication networks.

Content-based applications: Examples can be music and video sharing with bulk sizes. The current cellular networks are facing a data content explosion. According to Cisco forecast, the data content in the cellular network increase at the rate of 131% per year from 2008 to 2012. Even by upgrading all the network to 4G, the capacity will still not able to accommodate the growth. A sensible way is to offload some of the popular contents to a selected small group of users and rely on these users to further propagate the contents [11]. Underground stations and trains usually highly populated and hence good locations for content dissemination [12].

Social-networking applications: Bluetooth can enable a series of social-networking applications. By using the Bluetooth scan as we have done in our experiment, we can explore who are our “familiar strangers” [13] and who are always following us. Integrating the Bluetooth ID to the popular social network applications (e.g. Facebook and RenRen) can enable us to socializing whom are really physically close with us.

Gaming applications: Riding on the train can be very boring or very fun. Proximity-based interaction games can make the journey much more interesting. Metropolitan commuters can compete will each other during their move with short range wireless radio as they are playing online Internet games. Simple ones can be chess games and complicated ones can even provide virtual reality and include the physical locations of the users into the hints of the games.

VI. RELATED WORK

In this section, we introduce previous research efforts that have motivated us to propose our new idea. Bluetooth [7]

enables many new emerging applications making the use of a single mobile telephone much more than a cellular phone itself, and attracts much of the research and industry interests. Regarding the Bluetooth network formation, the Bluetooth Special Interest Group (SIG), unfortunately, has no specifications this far on how to form a scatternet [10], however in [8], the enhanced mobility aware routing (EMOLAR) is proposed to deal with the mobility of the devices by the scatternet formation and route reconstruction algorithms. Sarkar et al. discussed the complexity of topology formation based on the known “master-slave” model, with a constraint on the number of slaves a master can support. By investigating the Bluetooth complex implementation, they proved Bluetooth was unlikely to be widely used in large ad hoc networks, but it was still possible to be used in *small* ones successfully and this gives us the fundamental motivation to build our own application on top of it. However, in the traditional “master-slave” model, the communication between slaves relies on the master as a transit, as a big challenge on the master node. Our goal is to envision a relatively small Bluetooth network with multiple masters elected as an ad-hoc manner, and all other Bluetooth devices are able to share information in a much casual style.

With the popularity of smartphones today, peer-to-peer file sharing and social networks has drawn much attention in both academia and industries, where a major trend foresees the social networks tend to be mobile and the information disseminations among users can be achieved by social participation [11]. In [3] authors proposed an efficient file acquiring/distribution method from its neighbors in the opportunistic networks. They implemented it by a smartphone Bluetooth interface and achieves up to 417KB per second datarate. Our work is also largely motivated by [2] that proposed a novel efficient network management framework to improve the quality of information (QoI), also the trade-off between the maintenance of the the smartphone energy.

Human mobility models were studied in [14], [15], and passenger behavior model in transportation hub was discussed in [16]. More recently a permanent network of Bluetooth-enabled poster sites was deployed on the London Underground [6] to exploit the business value of the underground scenario where a survey has shown that passengers are open to advertisement in such environments. However, all these research efforts have no clear notion to consider how to best explore the passenger dynamics in an underground environment of major metropolitan cities like Beijing, and its inherited passenger behaviors and more importantly in which way communications and networking can be improved and leveraged. To this end, exploring the human factors, combined with the mobile social networks via Bluetooth are those of the research gaps; and these primarily drive our efforts in this paper.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we presented our recent experimental study on passenger dynamics and connectivities in Beijing Underground during the peak hours, via building an information collection application in an Android-based smartphone. From the total of

41,806 records for 120 days, we show the ECDF of a variety of statistics like the number of neighbors, the lifetime of the neighbors, and the RSSI and its derived physical distance, flow speed of the nearby passengers, etc. By using the results, we showed the Bluetooth’s applicability of building up a relatively small multi-hop wireless network with the network size of four nodes in such underground environments during the peak hour. We also presented a few interesting applications building on top of this Bluetooth-enabled multi-hop wireless networks.

In the future work, we are planning of building a testing prototype using Android phones as mobile clients and empirically evaluate the formation of such network and the transfer some files over it in a dynamic environment. We are also working on how to include more the social aspects like passenger preferences into the applications and redesign some more interesting applications. We believe our work is a fundamental stone for such emerging type of applications, or mobile Internet in general; and a lot of work can follow.

REFERENCES

- [1] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn, “The rise of people-centric sensing,” in *IEEE Internet Comp. Special Issue on Sensor Networks*, 2008.
- [2] C. H. Liu, P. Hui, J. W. Branch, C. Bisdikian, and B. Yang, “Efficient network management for context-aware participatory sensing,” in *IEEE SECON’11*, June 2011, pp. 116–124.
- [3] S. Lu, G. Chavan, Y. Liu, and Y. Liu, “Design and analysis of a mobile file sharing system for opportunistic networks,” in *ICCCN’11*, 2011, pp. 1–6.
- [4] S. Gaonkar, J. Li, R. R. Choudhury, L. Cox, and A. Schmidt, “Microblog: sharing and querying content through mobile phones and social participation,” in *ACM MobiSys’08*, Breckenridge, CO, USA, 2008, pp. 174–186.
- [5] N. Eagle and A. Pentland, “Social serendipity: mobilizing social software,” *IEEE Pervasive Comp.*, vol. 4, no. 2, pp. 28–34, 2005.
- [6] “Bluetooth network in london underground,” <http://www.cbsoutdoorinternational.com/Press-Centre/Press-Releases/Viacom-Outdoor-launches-Bluetooth-Network-on-London-Underground/>.
- [7] B. A. Miller and C. Bisdikian, *Bluetooth Revealed: The Insider’s Guide to an Open Specification for Global Wireless Communications / Edition 2*. Prentice Hall Professional, July 2002.
- [8] R. Debnath, B. C. Kundu, and M. Pradhanand, “Lossless secure transmission in bluetooth scatternet, considering device mobility,” in *2nd Int’l Conf. on Wireless VITAE*, March 3 2011, pp. 1–5.
- [9] T. M. T. Do and D. Gatica-Perez, “Groupus: Smartphone proximity data and human interaction type mining,” in *5th Annual Int’l Symp. on Wearable Computers (ISWC’11)*, June 2011, pp. 21–28.
- [10] “Core specifications,” <https://www.bluetooth.org>.
- [11] B. Han, P. Hui, V. Kumar, M. Marathe, J. Shao, and A. Srinivasan, “Occupancy grids: a stochastic spatial representation for active robot perception,” *IEEE Trans. on Mobile Comp.*, 2011.
- [12] C. M. L. McNamara and L. Capra, “Media sharing based on collocation prediction in urban transport,” in *ACM MobiCom’08*, 2008, pp. 58–69.
- [13] E. Paulos and E. Goodman, “The familiar stranger: anxiety, comfort, and play in public places,” in *ACM CHI’04*, 2004, pp. 223–230.
- [14] A. Mei and J. Stefa, “SWIM: A simple model to generate small mobile worlds,” in *IEEE INFOCOM’09*, April.
- [15] K. Farrahi and D. Gatica-Perez, “Probabilistic mining of socio-geographic routines from mobile phone data,” *IEEE J. of Sel. Topics in Sig. Processing*, vol. 4, no. 4, pp. 746–755, Aug. 2010.
- [16] Z. Qi, H. Baoming, and L. Fang, “Simulation model of passenger behavior in transport hubs,” in *ICIMA’09*, May 2009, pp. 220–224.