MAKING SENSE OF PROXIMITY-BASED PATTERNS IN A PUBLIC SPACE

Shin'ichi Konomi[1], Mingmei Li[1], Xiaowei Shao[1], Ryohei Suzuki[2]
Kaoru Sezaki[1,2], Ryosuke Shibasaki[1,2], Yuri Nakagawa[3]

Abstract
Integrating ad hoc wireless communication into dynamic, complex spaces is an inherently challenging task. In this paper, we analyze over 2,800 anonymous pedestrian trajectories captured in a busy train station so as to uncover the unique patterns that emerge as people walk and change mutual distances. We then discuss how our analysis of proximity-based patterns informs the design of communication infrastructures and social applications for pedestrians.

1. Introduction

People use short-range wireless communication technologies such as Bluetooth, Infrared, Wi-Fi, RFID and NFC (Near Field Communication) in street corners, coffee shops, public transport, convenience stores, and so on. Mobile phones and game consoles often exploit these technologies to provide the P2P communication services that allow colocated people to exchange phone numbers, text messages, photos, movie clips, virtual pets, and even digital cash. The use of these technologies in social events[4], emergency communication, and mundane urban life[8] has received much attention from researchers in recent years. Relevant research topics include mobile peer-to-peer applications for augmenting face-to-face interactions[5] as well as simulation-based experiments[6] that can expand the potential of the *ad hoc* networks among pedestrians. What seems to limit the impact of these types of research is the difficulty to understand the complex, collective patterns that emerge from the pervasive usage of mobile wireless devices in the real world. Recently, researchers examined Wi-Fi and Bluetooth traces to discuss the impact of a temporal characteristic on the design of opportunistic forwarding algorithms[2]. In this paper, we take a close look at human foot-movement data to understand the impact of various distances on the structure of proximity relationships. The revealed patterns suggest the need of careful design of communication infrastructures and social applications.

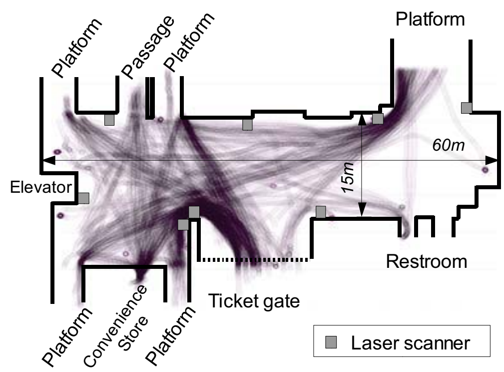


Figure 1: Pedestrian trajectories in *D*₁

2. The Pedestrian Trajectory Data

We use two datasets that were captured at a train station (see Figure 1) using 8 laser range scanners, which

[1] Center for Spatial Information Science, Univ. of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan
[2] Institute of Industrial Science, Univ. of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
[3] Research and Development Center of JR East Group, East Japan Railway Company, 2-0, Nisshin-cho, Saitama-shi, Saitama 331-8513, Japan

measure distances of nearby objects by emitting eye-safe laser beams at controlled directions and computing their time of flight[9]. The first dataset (D_1) was captured during the period between 7:00 a.m. and 7:05 a.m., and the second (D_2) 8:00 a.m. and 8:10 a.m. on a weekday. In the scope of this research, we explore as many characteristics as possible of the data that are not prohibitively expensive to process. In dataset D_1 and D_2, we found 367 and 2,438 pedestrian trajectories whose tracking accuracy is 93.4% and 88.7%, respectively. Their average time duration is about 20 seconds. In Figure 1, each semi-transparent line represents a pedestrian trajectory. Dark regions suggest crowded "hot spots," and people likely paused at places indicated by dark dots. Figure 2 shows temporal patterns in D_1 and D_2. It is more crowded around 8 a.m. than 7 a.m., and the rhythm of the place seemingly corresponds to arrivals and departures of trains.

The datasets include a massive amount of foot-movement data at 50 millisecond intervals. Calculating mutual distances of all human feet at each time point, we connect the pairs that are within the threshold distances of *1.2, 3.0* and *10.0* meters, which closely match Hall's personal, social, and public distances[3]. We then aggregate and visualize the *proximity networks* using the *neato* layout heuristic. In this layout heuristic, an ideal spring is placed between every pair of nodes and the length of the spring reflects the shortest path distance between the endpoints [7].

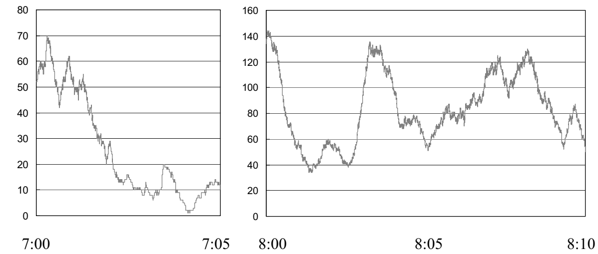
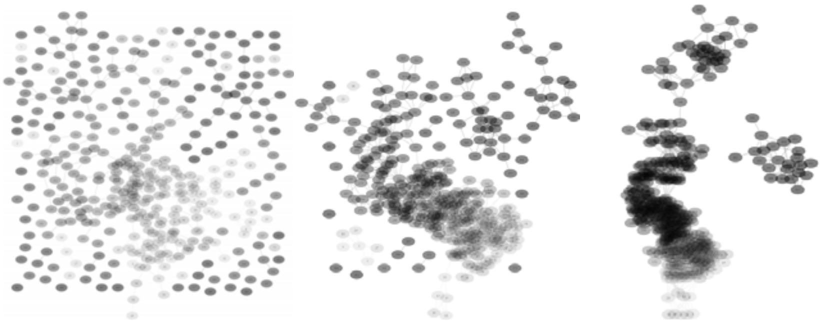


Figure 2: Number of people at the train station

3. Proximity-Based Patterns

As shown in Figure 3, the *neato* layout heuristic generates unique shapes that resemble a "winding trail" as we increase the threshold distance from *1.2m* to *10m*. The earlier pedestrians appear in the train station, the more transparent their nodes are in the network.

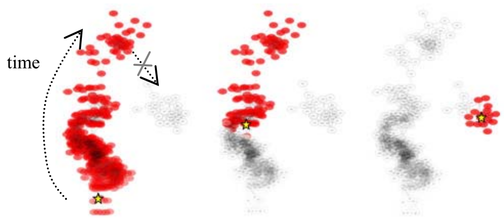


(a) *1.2m* (b) *3m* (c) *10m*

Figure 3: Proximity networks in D_1, with different threshold distances

Figure 4 shows how information can spread in the network of Figure 3c when it is sent from the 1st pedestrian (p_1), the 270th pedestrian (p_{270}), and the 355th pedestrian (p_{355}). The red nodes represent the pedestrians who can receive the information (through a series of wireless P2P communication.) p_1 walked towards the upper-right platform surrounded by the same group of people. p_{270} walked from the lower-left platform to the ticket gate passing by various people. p_{355} walked side by side with another pedestrian, from the ticket gate to the upper left platform. As shown in Figure 5, p_1 can reach everyone in the large cluster but the farthest people are 15 hops away.

These proximity networks connect any people who are in proximity; therefore a connected pair does not necessarily have a social tie. We nevertheless believe that it is meaningful to analyze these networks since they



(a) p_1 *(b) p_{270}* *(c) p_{355}*

Figure 4: How information spreads from the 1st, 270th, & 355th pedestrians in the network of Figure 3c

approximate the paths through which information flows via wireless devices. In Figure 6a, mean *closeness*[4] increases rapidly at *8m* and *33m*. Major clusters merged at these distances, making it much easier for a pedestrian to reach others. Mean *betweenness*[5] rapidly increases at the same distances; however, its standard deviation is more increasing than mean. This suggests that the "power" of pedestrians in the proximity network is less evenly distributed when the threshold distance is large. D_2 reveals much more moderate lines of closeness and betweenness than D_1. We speculate that it could be because it was very crowded around 8 a.m. possibly making people's distribution pattern in the space more uniform.

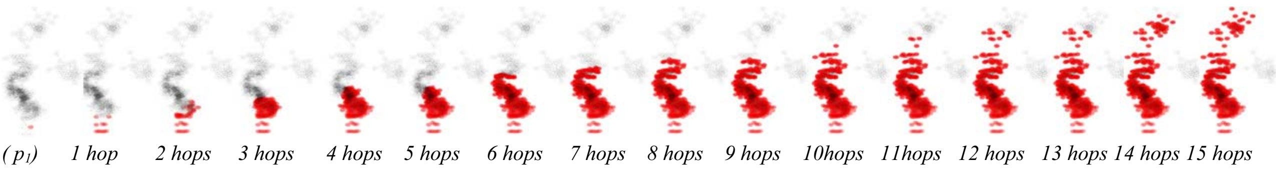


Figure 5: Hop-by-hop analysis of the message propagation from pedestrian p_1, in the network of Figure 3c

4. Discussion

Based on the analysis, we discuss the need of realistic mobility models, the impact of distances, and opportunities and risks in proximity-based communication environments.

Most mobility models [5] based on random waypoint directions are no longer adequate for understanding the environment constraints in the real world. From Figure 1, we notice that the pedestrian movement directions are not purely random. Most people follow the obvious and explicit orientation, such as the directions to some "hot spots" (e.g. a ticket gate); not surprisingly, the directions to some "cold spots" are very few. Therefore, we expect to provide a set of models that can be used to follow the real people movement directions. To understand other mobility characters from the pedestrian experiences, we also expect to extract the distributions from the data of the users' speed and pause time.

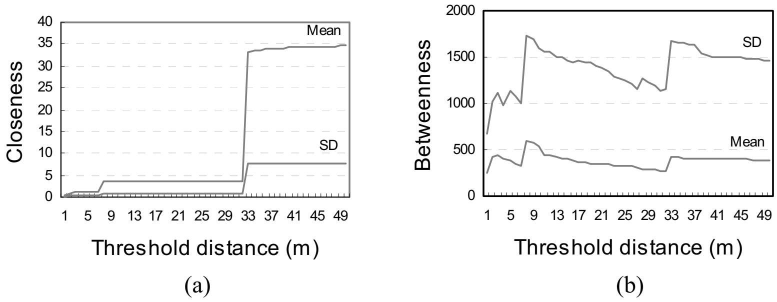


Figure 6: Analysis of network centrality for dataset D_1

Figure 6 suggests that we might have to be careful about the selection of a threshold distance. At this point, however, we don't have a conclusive answer as to whether or not similar staircase patterns exist in all public spaces. Not surprisingly, the figure also shows that one can reach others more easily by extending the spatial communication range of her device. Still, some people can be many hops away as suggested by Figure 5. Buffers and queues may have unique roles here as they can extend the communication range temporally. Information of course can be buffered digitally (e.g., read/write RFID tags) as well as physically (e.g., restrooms and convenience stores). Some of

---

[4] In essence, *closeness* is the reciprocal of *farness* which is calculated for any given node as the sum of the distances from the node to all other nodes.

[5] Betweenness is calculated for any given node as the proportion of shortest paths between all pairs of nodes that include the node.

the pedestrians with highest centrality scores are the ones who wandered about in the train station, stood still for a little while, and/or walked through the hot spots.

The question remains if the ultimate goal of the proximity-based wireless infrastructures is to connect any kinds of devices anywhere at anytime. We must simultaneously consider privacy, security, battery life and other issues as well as application scenarios. Relevant applications include emergency communication systems as well as locative media and social networking. In emergency communication, we need to find out the information transmission routes that allow people in some areas to keep in touch; to identify those in need of emergency assistance, and to share ideas on how to cooperate. We can find the possible transmission routes once the popular paths and connections are identified. Therefore, with the help of the pedestrian data, we expect to establish the emergency communication channels and improve the communication probabilities.

The anonymous, unlinkable laser scanner datasets were nevertheless highly useful for the purpose of analyzing proximity-based patterns. Although laser scanning systems are different from vision and RFID-based surveillance systems that could easily identify people, we must carefully consider social implications of instrumenting public spaces with sensing devices since it is extremely difficult to solve social issues such as privacy by retrofitting existing pervasive technologies.

5. Conclusion

We analyzed proximity-based patterns using the real-world data and discussed the impact of the physical context and mobility patterns on mobile ad hoc communication spaces. The data revealed the unique shapes and complex relationships, suggesting a need of careful infrastructure/ application design.

It is important for the users and designers of dynamic, complex pervasive systems to make sense of the real world patterns. Such sensemaking allows us to not only fit the technologies to existing practices but also create novel application environments and appropriate the technologies in meaningful ways.

6. References

[1] Borgatti, S.P., M.G.Everett, & L.C.Freeman (2002) Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

[2] Chaintreau, A., P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott (2006) Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms. *Proc. IEEE INFOCOM 2006*, pp.1-13.

[3] Hall, E.T. (1966) *The Hidden Dimension*, Doubleday.

[4] Holmquist, L.E., J.Falk, & J. Wigstrom (1999) Supporting Group Collaboration with Interpersonal Awareness Devices. *Personal Technologies 3*: 13-21.

[5] Kortuem, G., J.Schneider, D.Preuitt, T.G.C.Thompson, S.Fickas, Z.Segall (2001) When Peer-to-Peer comes Face-to-Face: Collaborative Peer-to-Peer Computing in Mobile Ad hoc Networks. *Proc. Int'l Conf. Peer-to-Peer Computing*, pp.75-91.

[6] Le Boudec, J.-Y. & M.Vojnovic (2005) Perfect Simulation and Stationarity of a Class of Mobility Models, *IEEE Infocom 2005*.

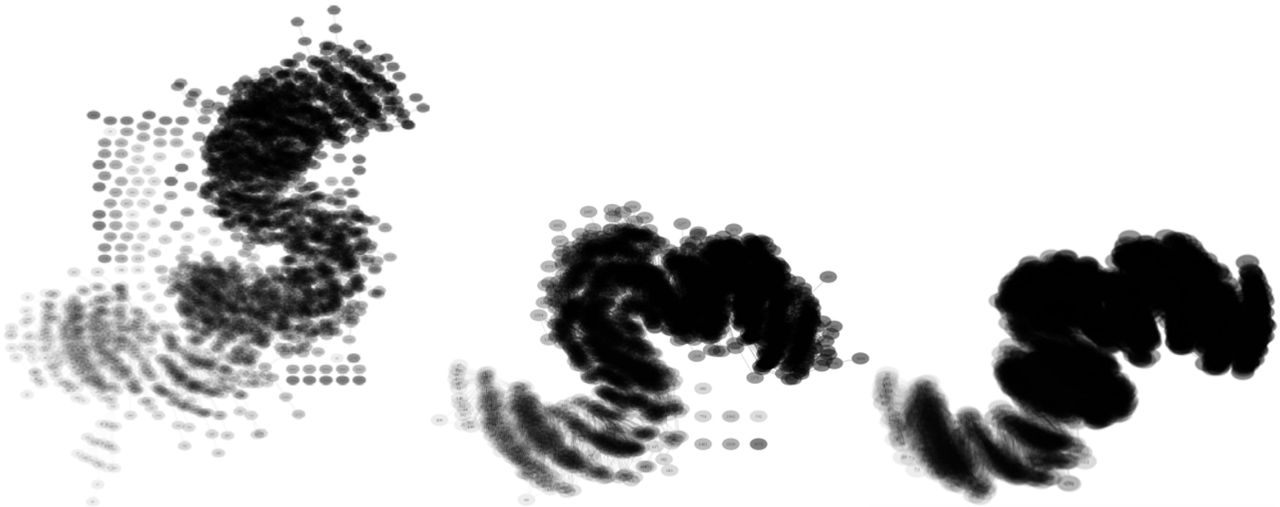
[7] North, S.C. (2004) Drawing graphs with NEATO. http://graphviz.org/Documentation/neatoguide.pdf

[8] Paulos, E., & E.Goodman (2004) The Familiar Stranger: Anxiety, Comfort, and Play in Public Places. *Proc. ACM CHI 2004*. pp. 223-230.

[9] Shao, X., H.Zhao, K.Nakamura, K.Katabira, R.Shibasaki & Y.Nakagawa (2007) Detection and Tracking of Multiple Pedestrians by Using Laser Range Scanners. *Proc. IEEE Int'l Conf. Intelligent Robots and Systems(IROS 2007)*, pp. 2174-2179.

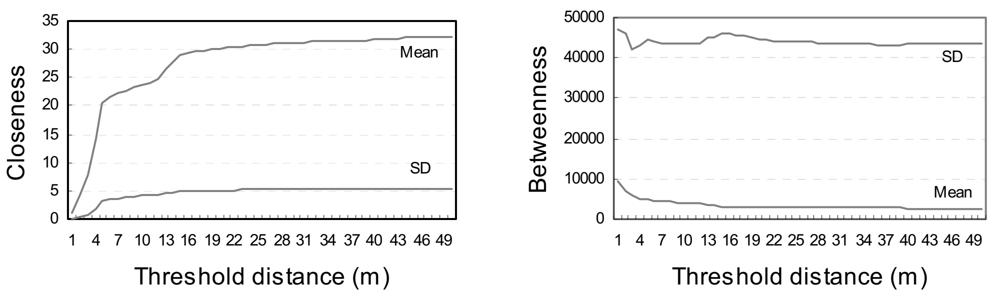
Appendix: Additional Information about the Poster

In our poster presentation, we will present the result of our analysis using the following images, graphs, and tables as well as the four figures included in the paper.



(a) *1.2m* (b) *3m* (c) *10m*

Proximity networks in dataset D_2 (8:00 – 8:10 a.m.), with different threshold distances



Analysis of network centrality for dataset D_2 (8:00 – 8:10 a.m.)

Dataset D_1 (7:00 – 7:05 a.m.)

	1.2m (personal distance)	3m (social distance)	10m (public distance)
Network density	*0.0075 (SD=0.0861)*	*0.0253 (SD=0.1570)*	*0.0870 (SD=0.2819)*
Number of cliques (>=3)	*112*	*582*	*1,834*
Degree centrality	*2.736(SD=2.566)*	*9.253(SD=6.397)*	*31.858(SD=18.428)*
Closeness centrality	*0.617(SD=0.184)*	*1.165(SD=0.386)*	*3.484(SD=0.845)*
Betweenness centrality	*418(SD=1,020)*	*442(SD=1,124)*	*535(SD=1,598)*

Dataset D_2 (8:00 – 8:10 a.m.)

	1.2m (personal distance)	3m (social distance)	10m (public distance)
Network density	*0.0028 (SD=0.0528)*	*0.0089 (SD=0.0941)*	*0.0303 (SD=0.1713)*
Number of cliques (>=3)	*2708*	*11,494*	*103,213*
Degree centrality	*6.824 (SD=4.649)*	*21.762 (SD=12.464)*	*73.736(SD=34.342)*
Closeness centrality	*1.060 (SD=0.075)*	*7.876(SD=0.593)*	*23.675(SD=4.174)*
Betweenness centrality	*9350.159(SD=46969.762)*	*5798.442(SD=42204.949)*	*4102.877(SD=43527.777)*

Network density, cliques, and centrality for the personal (1.2m), social (3m) and public (10m) distances