

Quantitative Evaluation of Public Spaces using Crowd Replication

**Samuli
Hemminki**

University of Helsinki
Helsinki, Finland
samuli.hemminki@
cs.helsinki.fi

**Keisuke
Kuribayashi**

University of Tokyo,
Tokyo, Japan
kuri@
csis.u-tokyo.ac.jp

**Shin'ichi
Konomi**

University of Tokyo,
Tokyo, Japan
konomi@
csis.u-tokyo.ac.jp

**Petteri
Nurmi**

University of Helsinki
Helsinki, Finland
petteri.nurmi@
cs.helsinki.fi

**Sasu
Tarkoma**

University of Helsinki
Helsinki, Finland
sasu.tarkoma@
cs.helsinki.fi

ABSTRACT

We propose crowd replication as a low-effort, easy to implement and cost-effective mechanism for quantifying the uses, activities, and sociability of public spaces. Crowd replication combines mobile sensing, direct observation, and mathematical modeling to enable resource efficient and accurate quantification of public spaces. The core idea behind crowd replication is to instrument the researcher investigating a public space with sensors embedded on commodity devices and to engage him/her into imitation of people using the space. By combining the collected sensor data with a direct observations and population model, individual sensor traces can be generalized to capture the behavior of a larger population. We validate the use of crowd replication as a data collection mechanism through a field study conducted within an exemplary metropolitan urban space. Results of our evaluation show that crowd replication accurately captures real human dynamics (0.914 correlation between indicators estimated from crowd replication and visual surveillance) and captures data that is representative of the behavior of people within the public space.

CCS Concepts

• **Human-centered computing**→Ubiquitous and mobile computing. • **Information systems**→Information systems applications→Spatial-temporal systems.

Keywords

Urban computing, crowd replication, mobile sensing, public spaces, evaluation

1. INTRODUCTION

Public spaces serve an important role in contemporary cities. Among others, they provide a medium for passive and active engagement and a place for discovery and reflection [2]. Public spaces have also been linked with quality of life indicators [5].

While benefits of public spaces have been extensively investigated, designers currently struggle with quantifying whether a particular space promotes the desired activities and benefits. Indeed, surveys

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

SIGSPATIAL'16, October 31-November 03, 2016, Burlingame, CA, USA

© 2016 ACM. ISBN 978-1-4503-4589-7/16/10 \$15.00

DOI: <http://dx.doi.org/10.1145/2996913.2996946>

and direct observation remain the dominant approaches for assessing public space quality [8][16]. In the former case, quality of different facets, ranging from psychological to physical and social variables, is assessed. However, these can seldom be directly correlated with behaviors or activities within the space. In the latter case, people and their behavior within the space are observed and indices related to, e.g., the amount of mobility and activities within the space are derived from the observations. A major limitation of this approach is that collecting data is labor intensive and time consuming, often resulting in a small sample size. While pervasive sensing technologies, e.g., based on visual surveillance [12], location tracking [14], crowd sensing [15][17][19], and analysis of cellular records [1], can support data collection, they have not been widely adopted due to high resource cost, the need for expertise in interpreting the data, and privacy concerns. Many of these techniques also induce unknown biases to the data, thus limiting generalizability of the conclusions that can be drawn.

The present paper contributes by developing crowd replication as a low-effort, easy-to-implement and cost-effective mechanism for quantifying the uses, activities, and sociability of public spaces. The core idea in crowd replication is to engage the researcher assessing the public space into replicating the behavior of people within the space. A commodity smartphone (or another smart device such as a smart watch) equipped with location and motion sensors (i.e., accelerometer and gyroscope) is carried by the researcher and used to simultaneously capture a sensor trace corresponding to the observed behaviors, thus effectively replicating the observed behavior. By combining the sensor traces with a population model, indicators reflecting the aggregated behavior of a large amount of people can be calculated from a small subset of observations. As the bulk of the data collection is based on mobile sensing, user involvement and overall effort are significantly reduced. Another benefit of crowd replication is that it mitigates privacy issues by considering only measurements that reflect the movements of the researcher. Accordingly, only anonymized data of individuals is revealed to designers.

We demonstrate the benefits of crowd replication by carrying out a case study of crowd replication within a representative example of a metropolitan-scale public space. As part of our case study, we compare the data quality and overall effort of using crowd replication against the use of video tracking. Our results show that crowd-replicated data closely mirrors (0.914 correlation between crowd replicated data and visual tracking data) the real human dynamics in public spaces and are useful building block for human-scale analytics.

2. QUANTIFYING USAGE OF PUBLIC SPACES USING CROWD REPLICATION

Figure 1 gives an overview of the crowd replication process. First, during *Study Preparation*, the target environment is investigated to identify important locations and other factors that influence sampling the active population of the target area. Next, during *Data Collection*, the researcher carrying out the investigation is tasked with following and observing people moving in the space, replicating their trajectory and activities on general level. The researcher is also equipped with a commodity smart device, which runs a mobile application that creates a sensor trace corresponding to the replicated behavior by recording location and motion measurements of the movements and behavior of the researcher. Finally, during *Analysis*, the collected sensor data is used to reconstruct location trajectories and activities of users. The individual traces are then combined with a population model that can be used to scale observations drawn from individual users to reflect the aggregate behavior of the entire target population in a representative and a generalizable manner.

2.1 Study Preparation

During *Study Preparation*, a sampling strategy for selecting which activities and people to replicate is chosen. The key requirement for sampling is to choose the participants and activities that are being replicated so that the resulting data is representative of the demographics and activities of the place. In our experiments we consider a strategy inspired by cluster sampling where the entry points of a public space are used to select people to replicate.

After choosing the appropriate sampling strategy, a mobile application responsible for recording sensor measurements is configured for the area of investigation. First a localization technique is selected for location tracking, e.g., pedestrian dead reckoning combined with GPS or indoor localization. Physical activities are detected using motion sensors (i.e., accelerometer and gyroscope) as such activities can typically be inferred with high confidence. The mobile device can also be configured to assist the researchers by automating the randomization associated with the sampling strategy. Additionally, the app can be used to add annotations relevant for the current user replication trace. Typical annotations include user or group demographics, social or other such activities that cannot be discreetly replicated, and a free field for notes on abnormalities occurring during the trace.

2.2 Data Collection

Next, during *Data Collection*, the researcher carrying out the investigation is tasked with following and observing people moving in the space, replicating their trajectory and activities on general level. The researcher is also equipped with a commodity smart device, which runs the mobile application that creates a sensor trace corresponding to the replicated behavior by recording location and motion measurements of the movements and behavior of the researcher. The aim is to produce a trace of sensor measurements which correspond to the motion of the target user as closely as possible, i.e., the trace would correspond to sensor measurements which the target person would have produced if she had been wearing the sensing devices. For most studies, the most important aspects to monitor are the trajectories and general physical activities of the users within the space. We capture these automatically through motion sensing and location inference techniques integrated onto an app that is running on the device carried out by the researcher.

Motion Sensing: Our motion sensing operates on three dimensional acceleration and angular velocity measurements. From acceleration

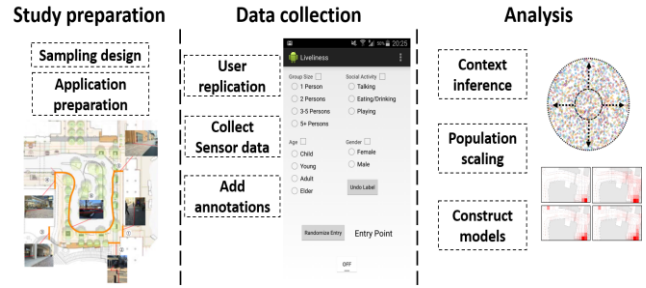


Figure 1. Overview of crowd replication process

measurements, we calculate the overall motion level by leveraging a well-established method of computing the standard deviation of the L_2 -norm [7]. Next, we derive the user’s walking pace by using the angular velocity measurements obtained from gyroscope. First, we compute autocorrelation of the measurements, after which we select the autocorrelation peak that minimizes the distance between consecutive cycles. As distance measure we use dynamic time warping (DTW). The intuition is that the first stage produces one or more candidates for the correct walking pace, while the second stage identifies the correct one amongst the candidates. As a post-processing technique, we verify that the detected walking pace agrees with a linear regression model between the overall motion intensity and walking pace. This post-processing step is particularly effective as errors in walking pace are typically multiples of the correct walking cycle. Consequently, these errors can be easily detected as a large deviation from typical relation between standard deviation and walking pace. The resulting walking pace estimation is accurate up to >99%, only failing in rare occasions.

Location Inference: Our location inference approach fuses localization with pedestrian dead reckoning (PDR) to produce a trajectory of the user’s movements. In our experiments, we consider an outdoor space and hence rely on GPS for localization. We sample GPS at 1Hz, which ensures high tracking accuracy while allowing sufficient battery lifetime to last through the study. GPS is used to recalibrate PDR and to reduce drift, whereas the use of PDR helps to increase location accuracy, particularly when GPS accuracy is low. Our PDR implementation is based on the walking pace detection module (see above), and a sensor fusion method proposed by Madgwick [10]. We balance between localization and PDR by weighting localization information based on an accuracy estimate x , e.g., horizontal accuracy reported by GPS. As the initial location of user’s motion trajectory, we utilize coordinates of the respective entry point.

2.3 Analysis

During *Analysis*, the collected sensor data is used to reconstruct location trajectories and activities of users. The trajectories are then combined with labels annotated by the researcher. These trajectories alone are sufficient for answering many questions regarding the use of public spaces as many characteristics are naturally quantified using relative values. As a few example cases, consider evaluation goals such as inspecting motion flows within the area, fraction of people stopping at an information post or exhibition stand, or ratio between young and elderly age groups. Many other characteristics, however, require values that reflect the total number of individuals. For assessing these characteristics, it is necessary to generalize the indicators derived from replication to the entire population of the target area. We accomplish this by deriving scaling factors that enable adjusting values from replicated data into the overall population.

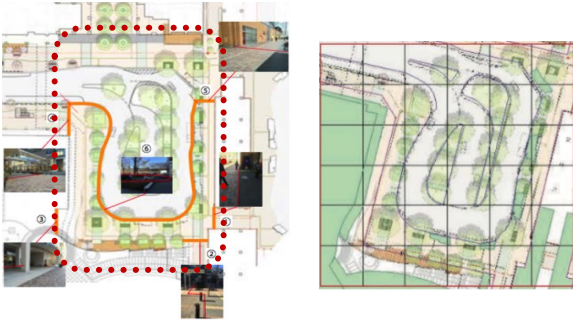


Figure 2. Center area and its main entry points (left), and the area split into 15x15m grids (right).

To formalize this process, let P denote the total population of the target area. This information can usually be approximated from national survey data, e.g., pedestrian count data [3] [4], residence counts, or travel survey data. Without loss of generality, we assume the replication process has been carried out using the cluster sampling, i.e., that the replicated data can be grouped into a discrete number of clusters. Let R_i denote the number of users replicated within cluster i , i.e., in our case the number of replicated users from entry point i and let E_i denote the frequency of users collected at the same entry point. For each cluster i we derive a scaling coefficient S_i using:

$$S_i = (P * E_i) / R_i \quad (1)$$

Let $f(D_i)$ denote the value of the indicator estimated from the replicated measurements D_i . The adjusted estimator $f(D')$ for the entire population can then be simply calculated using

$$f(D') = \sum_{i=1..n} S_i * f(D_i) \quad (2)$$

Assuming the set of trajectories replicated in the study is representative of the behaviors of the people within the space, the scaling coefficients enable generalizing indicators derived from the replicated data to the entire target population.

The scaling procedure described above only takes into consideration active population, i.e., users who are moving or performing activities of interest. In certain cases, we are also interested in measures that take into account users that are stationary or that are performing activities that cannot be reliably replicated (i.e., which are captured using manual annotation only). Crowd replication directly generalizes to these scenarios by offering a method to consider subpopulations which were not subjected to replication. Specifically, assume we are interested in a subpopulation s . Let f_s denote the fraction of the population belonging to subpopulation s and let R_s denote number of samples from the subpopulation s . We can now derive a scaling coefficient S_s and scaled indicators $f(D_s')$ of the subpopulation data using the following equations:

$$S_s = (P * f_s) / R_s \quad (3)$$

$$f(D_s') = S_s * f(D_s) \quad (4)$$

Finally, to mitigate bias due to varying lengths of trajectories, we normalize the scaling factors with respect to the length of the trajectories, i.e., we derive a normalized estimator $f(D_{norm})$ using:

$$f(D_{norm}) = l * \sum_{i=1..m} d_i / l_i \quad (5)$$



Figure 3. Target public space and its nine entry points.

Here d_i is the indicator drawn from trajectory i and l_i is the length of i -th trajectory, and l is the mean length of the trajectories.

2.4 Discussion: Privacy and Ethical concerns

Recording only sensor measurements instead of other mediums, e.g., video or audio information [12], has the added benefit of increasing privacy of the tracked users. The primary advantage is that this information cannot be used to identify users. However, we acknowledge that the process of user behavior replication should be carried out in a way that avoids any privacy or ethical concerns. The necessary constraints depend on the target space, but should be planned so that the user behavior replication can be conducted without the target user becoming aware or disturbed by the replication. This is also crucial for reducing possible observer biases. As an example, we kept 10-15 meters distance between the researcher and the target and aborted the process in case the target seemed to become aware of the researcher.

3. EMPIRICAL EVALUATION

We next validate the use of crowd replication for data collection by comparing dwell time indicators extracted from crowd replication with those extracted from video surveillance. In the following we first describe our target area, after which we present the results of our evaluation.

3.1 Target Area

We have conducted a field study at Kashiwa-no-ha smart city, portrayed in Figure 2 and 3. The area is located within the larger metropolitan area of Tokyo and is at the intersection of a popular shopping center and transportation hub encompassing train, bus and taxi stations. We selected this area since it includes various artifacts and spaces that have been designed to improve human experiences. The planners and designers are keen to evaluate the area quantitatively and recently conducted direct observation and manual pedestrian counts at several locations within the area. The area is mainly aimed for pedestrians, with cycling prohibited within the area. The area features services such as several sets of tables, total of 95 chairs and benches, information stand, guide map and an interactive touch display. The center of the area, as portrayed in Figure 2, contains six main entry points, and the full area, portrayed in Figure 3, contains nine entry points. As is typically the case with similar physical environments, the space is the product of a careful design by urbanists who desire to evaluate their design by measuring its usage in detail.

3.2 Validity of Crowd Replication

We validate crowd replication by comparing the distribution of motion flows against a ground truth obtained from video recordings. The video was recorded simultaneously with the crowd

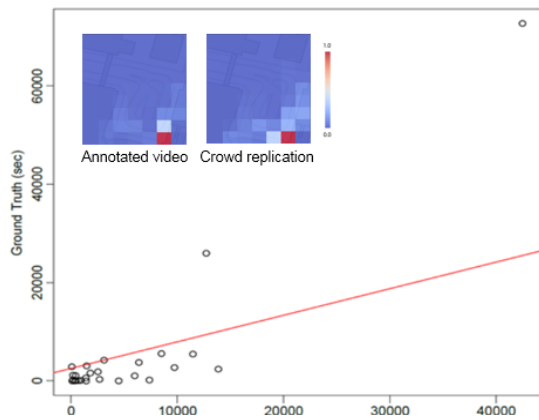


Figure 4. Comparison of normalized aggregated dwell time between manually annotated video and crowd replication, and comparison of spatial distribution of aggregated dwell time.

replication process from 6th floor balcony of an adjacent building. The distribution of activities is defined by separating the target area into 15m x 15m grid cells (see Fig. 4) and calculating *aggregated dwell time* within each grid cell. Aggregated dwell time is calculated as the total sum of duration over all the people passing through the area. Thus, lingering people increase dwell time more than mere passers-by, and it is a basic component for quantifying the liveliness of public spaces [11].

The results of this evaluation, presented in Figure 5, show a strong positive correlation between crowd replication and the video-based ground truth $r = .914$, $p < .01$. The remaining location errors are mainly due to limitations with localization accuracy, which in our case are emphasized by an area enclosed by tall buildings. We also note that the distribution of user demographics were similar for all researchers performing crowd replication, indicating that crowd replication is easy to implement in practice and capable of effectively capturing the actual distribution of users.

4. SUMMARY AND DISCUSSION

We contributed by developing crowd replication as a novel and cost-effective mechanisms for collecting data from public spaces. Experimental evaluations carried out within a metropolitan scale urban space demonstrated that the data collected with crowd replication closely mirrors the actual behavior of crowds within public spaces. As crowd replication allows sampling from all people in a public space, regardless of their ownership of mobile phones or other sensing devices, it has minimal bias. By combining a small amount of sensor traces with a population model, reliable indicators reflecting the aggregated behavior of a large amount of people can be obtained. Compared to opportunistic sensing which focuses on the scales of cities and greater regions, crowd replication can collect much finer-grained data in terms of spatial resolution and human activities, as well as discover subtle patterns of human motion. An additional benefit of crowd replication is that it can cover a relatively large area that encompasses streets and neighborhoods, which would be too large to cover by using conventional direct observation and visual surveillance. Another benefit from crowd replication is that it overcomes potential privacy issues by considering only anonymized measurements that reflect the movements and activities of the researcher. Our trials took place at outdoor sidewalks near a train station. The crowd replication method itself could be adopted for other public spaces, including malls, airports and parks, for different purposes, including comparisons across different demographic attributes and evaluations focusing on baby carriages or tourists, and for different

cultural contexts and social norms. Doing so can involve uses of different sampling rates and strategies and indoor location tracking technologies.

ACKNOWLEDGMENTS

The work was supported by Nokia foundation scholarship, by the Finnish Academy of Science under the project *Sampling in Pervasive Sensing Systems*, CREST, JST, and MEXT under the Green Network of Excellence (GRENE) program. We thank the members of Spatial Planning and Design Laboratory at the University of Tokyo for sharing the pedestrian count data. The work only reflects the authors' views.

REFERENCES

- [1] Calabrese, F., Ferrari, L., and Blondel, V.D. Urban Sensing Using Mobile Phone Network Data: A Survey of Research. *ACM Computing Surveys* 47, 2 (2014), 1–20.
- [2] Carr, S., Francis, M., Rivlin, L., Stone, M., Public Space, Cambridge University Press, 1993
- [3] City Of Melbourne, Pedestrian Counting System, <http://www.pedestrian.melbourne.vic.gov.au>, retrieved 29th Jun 2016.
- [4] City of New York, Times Square pedestrian counts | NYC Open Data, <https://data.cityofnewyork.us/City-Government/Times-Square-pedestrian-counts/dv6z-emb2>, retrieved 29th Jun 2016.
- [5] Spatial Planning and Design Laboratory at the University of Tokyo, <http://udcx.k.u-tokyo.ac.jp/Top.html>, retrieved 29th Jun 2016.
- [6] Gehl, J. *Life Between Buildings: Using Public Space*. Van Nostrand Reinhold, New York, 1987.
- [7] Hemminki, S.; Nurmi, P. & Tarkoma, S. Accelerometer-Based Transportation Mode Detection on Smartphones Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys), ACM, 2013.
- [8] Jacobs, J. *The Death and Life of Great American Cities*. Random House, New York, 1961.
- [9] Lazer, D., Kennedy, R., King, G., Vespignani, A., *The Parable of Google Flu: Traps in Big Data Analysis*, Science, 343, 2014, pp. 1203 – 1205
- [10] Madgwick, S. 2010. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. Department of Mechanical Engineering, University of Bristol, April 2010.
- [11] Mehta, V. Lively Streets: Determining Environmental Characteristics to Support Social Behavior. *Journal of Planning Education and Research* 27, 2 (2007), 165–187.
- [12] Placemeter, I. Placemeter, <http://www.placemeter.com>
- [13] Project for Public Spaces. *How to Turn a Place Around*. Project for Public Spaces, Inc., New York, 2000.
- [14] Quuppa, <http://quuppa.com/>, retrieved 29th Jun 2016.
- [15] Schauer, L., Werner, M., Marcus, P., Estimating crowd densities and pedestrian flows using wi-fi and Bluetooth, *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, (MOBIQUITOUS)*, 2014.
- [16] Whyte, W.H. *The Social Life of Small Urban Spaces*. The Conservation Foundation, Washington, D.C., 1980.
- [17] Walkbase indoor positioning platform, <http://www.walkbase.com/>, retrieved 29th Jun 2016.
- [18] Wang, J-F., Stein, A., Gao, B-B., Ge, Y, *A review of spatial sampling*, Spatial Statistics, 2, 2012, pp. 1-14.
- [19] Weppner J., Lukowicz, P., Blanke, U., and Tröster, G., Participatory Bluetooth Scans Serving as Urban Crowd Probes, IEEE Sensors, 2014.