Denne artikel er udgivet i det elektroniske tidsskrift
Artikler fra Trafikdage på Aalborg Universitet
(Proceedings from the Annual Transport Conference
at Aalborg University)
ISSN 1603-9696
https://journals.aau.dk/index.php/td

Context-Aware Sensing and Implicit Ground Truth Collection: Building a Foundation for Event Triggered Surveys on Autonomous Shuttles

Valentino Servizi, valse@dtu.dk
Department of Technology, Management and Economics, Technical University of Denmark (DTU)

Dan Roland Persson, Per Bækgaard
Department of Applied Mathematics and Computer Science, Cognitive Systems, DTU

Hannah Villadsen
Department of People and Technology, Roskilde University

Inon Peled , Jeppe Rich, Francisco C. Pereira, Otto A. Nielsen
Department of Technology, Management and Economics, DTU

This project is co-financed by the European Regional Development Fund through the Urban Innovative Actions Initiative.

Abstract
The LINC project aims to study interactions between passengers and autonomous vehicles in natural settings at the campus of Technical University of Denmark. To leverage the potential of IoT components in smartphone-based surveying, a system to identify specific spatial, temporal and occupancy contexts relevant for passengers' experience was proposed as a central data collection strategy in the LINC project. Based on predefined contextual triggers specific questionnaires can be distributed to affected passengers. This work focuses on the data-based discrimination between two fundamental contexts for LINC passengers: be-in and be-out (BIBO) of the vehicle. We present empirical evidence that Bluetooth-low-energy beacons (BLE) have the potential for BIBO independent classification. We compare BLE with other smartphone onboard sensors, such as the global positioning system (GPS) and the accelerometer through: (i) random-forest (RF); (ii) multi-layer perceptron (MLP); and (iii) smartphone native off-the-shelve classifiers. We also perform a sensitivity analysis regarding the impact that faulty BIBO ground-truth has on the performance of the supervised classifiers (i) and (ii). Results show that BLE and GPS could allow reciprocal validation for BIBO passengers' status. This potential might lift passengers from providing any further validation. We describe the smartphone-sensing platform deployed to gather the dataset used in this work, which involves passengers and autonomous vehicles in a realistic setting.

Introduction

A fleet of autonomous shuttles received approval for public operations at the Technical University of Denmark, starting April 2021. The trial is conducted as a part of the UIA funded LINC project with the purpose of studying future scenarios involving shared autonomous vehicles and their impact on society. To collect relevant user data and the information necessary to successfully pursue the project goals, LINC deployed a context-aware smartphone-based travel survey (SBTS) integrated with the autonomous vehicles and other Internet of Things (IoT) components—(i.e.) Bus Stops and Buildings—through Bluetooth low energy beacons (BLE) technology [1].

The implications of introducing context-aware smartphone-based travel surveys (SBTS) in the field of shared autonomous travel are very broad. Detection of spatial and temporal occupancy contexts can be leveraged to submit specific surveys to users who have had specific types of exposure to the service which can potentially reduce known problems with hypothetical-, recall-, and status-quo-bias in user data [2, 3]. Some examples of specific areas where this method provides opportunities to advance understanding of users travel behavior are: (i) the potential of improving public service through responsive time schedule; (ii) the effects of driverless shuttle service on social inclusiveness [4], perceived safety and security in specific contexts [5, 3] (e.g. darkness, presence of other passengers, unexpected driving patterns, emergency stops); and (iii) the impact on user's future transportation mode and route choices [3, 6].

To monitor passengers' presence aboard a vehicle—representing the basic context of interest for LINC common solutions—require users' explicit interactions with a device, e.g, to check-in/check-out (CICO), or the implicit interaction with a device detecting, e.g., walk-in/walk-out (WIWO). This work focuses on the implicit interaction between BLE and passengers' smartphones, which we identify with be-in/be-out (BIBO) [7].

The study was conducted during the development phases preceding LINC operations. We describe the architecture of the smartphone sensing platform and the results of the experiments directed to assess BIBO detection with BLE and Global Positioning System (GPS) data, with respect to smartphone native activity recognition based on, e.g., accelerometer. The empirical validation of this architecture involved passengers and vehicles in a semi-controlled environment, set up to approximate a highly congested urban area. Passengers were free either to walk or to travel on the autonomous vehicles. The environment included three bus stops and two buses operating on two routes, which intersected in one of the stops.

Results show that a BIBO system based on GPS and BLE allow higher accuracy than the native activity classifier distributed with Apple and Google operating systems, based on accelerometer, gyroscope and pedometer [8].

Related Work

The literature has shown several aspects that underpin our work. BIBO systems based on BLE protocol, which branches from the WiFi protocol, leverages properties such as frequent broadcast communication of a few bits of data and requires no paring between devices. These properties are desirable for proximity detection [9], and some proofs of concept have shown the potential of BIBO systems [10, 1].

Literature covering indoor smartphone tracking systems based on BLE is extensive [11]. Some of the main methods deployed are based on the Friis equation [12], trilateration [13], and fingerprinting [14, 15, 16]. The first two methods rely on the known position of the BLE devices; the last, on machine learning (ML) methods.

Bluetooth is considered an unpromising technology for Vehicle to anything (V2X) communication in intelligent transport systems due to the slow paring process [17]. However, BLE can trigger events on smartphones without any paring operations [18]; for example, the proximity of a smartphone to a BLE device could start an application and trigger events.

Outdoor tracking systems based on BLE and WiFi signals [19] allow transport mode detection based on received signal strength (RSSI) and ML methods, such as Random Forests (RF).

The pervasive use of supervised ML methods with smartphone signals in general [9], and with BLE in particular, exposes the potential impact of label noise on the training process of a ML classifier. The study of this problem counts on a vast literature focusing on independent and identically distributed (IID) datasets [20, 21, 22, 23]. However, we have found no examples focusing on time series signals, such as GPS, and BLE.

From the literature, the main takeaways for a successful implementation of a BIBO system based on BLE are the following: (i) BLE signal transmission rate above 0.3 Hz; (ii) The density of the BLE beacons network above one device every 30 square meters; and (iii) Appropriate imputation of RSSI readings.

Methods and Materials

This section presents the sensing platform, consisting of the smartphone application used for data collection, the experiment setup and the BIBO sensitivity analysis method.

Sensing Platform Overview

The data collection platform consists of several major components (see Fig. 1), firstly native smartphone applications handling system driven data collection on a per participant level. Secondly, several web services APIs providing live data from the autonomous shuttles and other sources. Thirdly, overlap between the two data sources enables simple but context-aware services to prompt users for user driven data to be collected in the form of surveys. The smartphone platform consists of two native apps for iOS and Android respectively coded in Swift and Java.

Each application collects the same data namely, location, sensor data, activity, and beacon proximity. The latter serving as a ground truth (GT) validation mechanism for collected data, providing a proximity measure to certain locations such as autonomous shuttles, buildings, and public transport nodes (see Fig. 2).

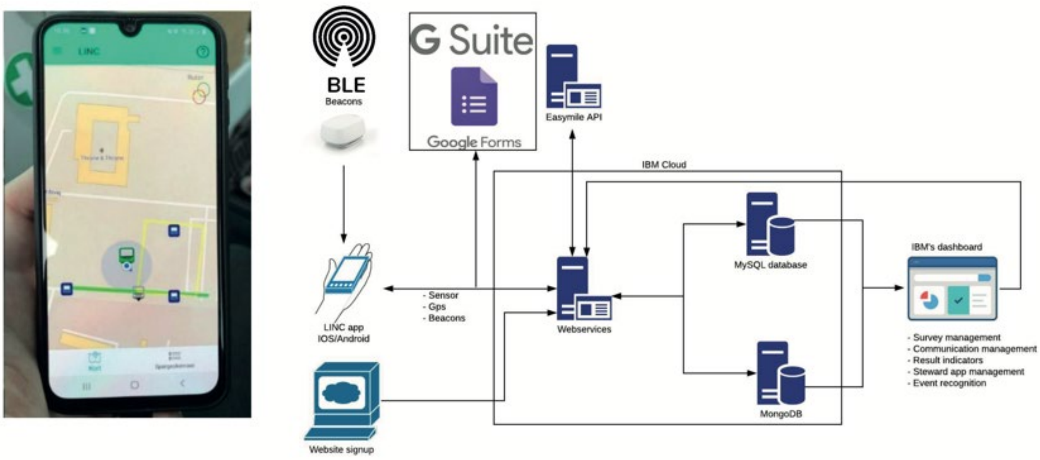

Figure 1 - Sensing platform

Additional data is collected from the autonomous vehicles using several web APIs, these include weather, vehicle status (including GPS, speed, weight, potential errors) and passenger counts. Using the smartphone and API data collection as sources, basic event recognition has been implemented to prompt user driven data collection in the form of ecological momentary assessment surveys for the specific events.

Figure 2 - BLE network - BLE beacons installed on bus, bus stop and building. The figure shows an example of GPS trajectory automatically labeled with BLE proximity.

Smartphone Application

The two native applications, handle collection of travel data automatically uploading the data to a webservice when possible, storing it locally if not. The collected data consists of 1) location, 2) sensor data, 3) activity recognition and 4) BLE beacons.

1) GPS data measured and recorded at 1Hz when the user is in motion, collection automatically enters sleep mode if no activity is recorded and is limited to a geofenced location of interest for the study.
2) Sensor data in the platform has been designed in a flexible manner allowing researchers the option of recording sensor data at a rate between 1-30Hz, these sensors being accelerometer, magnetometer, gyroscope.
3) Activity recognition is handled by native implementations determining the activity carried out by the user, such as: driving, walking, being stationary, or being on a bike. Activity is recorded at 1Hz.
4) BLE beacons are similarly recorded at 1Hz rate, providing beacons ids, general information and proximity measurements to a given beacon. As the project's beacon deployments are at well-known geographic locations these provide a secondary measure for proximity to known locations.

In addition to the smartphone platforms data collection abilities, the application also provides information related to the study, such as information about the autonomous vehicles (schedules, routes, disruptions, and news). Lastly, the app facilitates the delivery of momentary ecological assessments in the form of surveys. The delivery of these surveys being determined by a simple event recognition system using data from the smartphone platform and other data sources.

Experimental Setup

The experiment involved three bus stops and two buses operating on two routes, which intersected in one of the stops. BLE devices were installed inside each bus and at each bus stop. The environment was under the surveillance of three high resolution cameras, one for each bus stop. The three cameras provided full coverage of the track, thus allowing the surveillance of passengers mounting and alighting of any bus and walking between bus stops. The footage from these cameras allowed collection of high-quality ground-truth. The experiment procedures consisted of providing participants with a written description of the experiment, which they should read, then sign for consent. After acceptance and a short briefing, participants accomplished the following steps: (i)

Install the smartphone application. (ii) Read general conditions for the smartphone platform and provide relevant permissions to the app. These include smartphone sensors, location, and an activity-classifier, performing transport mode detection. (iii) Participants equipped themselves with 2 ID stickers worn on the front and back, to ease the ground-truth collection from video recordings. (iv) Participants start using the shuttles with the goal of completing a tour back to the starting point, mounting and alighting the bus in any order, with the possibility of walking between bus stops.

In total 18 participants completed the experiment, 2/3 of which allowed the collection of 13, 723 data points and high-quality ground truth on BIBO binary labels, each point representing approximately 1 second of their journey. A total of 6 participants did not provide the correct app permissions resulting in no smartphone data being collected causing a high number of unavailable trajectories.

BIBO Sensitivity Analysis

To assess BLE signal as independent measure for a BIBO system, we trained and evaluated under label-noise conditions two supervised ML classifiers: RF and MLP [24]. The process involved independently BLE RSSI and GPS signals for both these classifiers. We used Monte Carlo simulations to assess the effect of various levels of noise causing labels' values to flip from BI to BO or vice-versa. Then we compared the performance of classifiers trained without flipped labels, against classifiers trained with flipped labels. As baseline, we referred to the native activity-classifier (see Sec. 3.2 and 3.3) distributed with Apple and Android operating systems (OS), where BI corresponds to "automotive", and BO to everything else [8].

A grid search using only a random subset of users - corresponding to approximately 80% of the dataset - and 5-fold validation scheme provided optimal hyper-parameters for training. The evaluation of each classifier trained with ground truth and optimal hyper-parameters consisted of repeating the following steps 100 times: (i) Pick a random number of users from the full dataset; (ii) Train the classifier with this partition, sized approximately 80% of the population; (iii) Test the classifier with the complementary partition.

To test the impact of noisy labels on classifiers training, one step preceded the evaluation process. In such a step a portion of the labels was flipped, simulating labeling errors. In multiple iterations, we tested a systematic increase of the proportion of flipped labels in the range from 0 to 100%. At the extremes of this range, where none or all the labels are flipped, the BIBO classification is equivalent. Area Under the Curve (AUC) [25] is the metric of reference to compare differing conditions, as explained in this section. AUC range is between 0 and 1, where 0.5 corresponds to the random classifier, or to a classifier predicting consistently only one out of two classes. AUC > 0.5 corresponds to a classifier predicting consistently the target class, the higher the better. AUC < 0.5 corresponds to a classifier predicting consistently the opposite of the target class.

Results

Results show that the difference between the random classifier and the baseline is minimal (see Fig. 3 and 4), and highlights that the BIBO classification task can be extremely challenging when people and vehicles move in proximity and at low speed, i.e., < 15km/h. On the opposite side of the scale, and in the same experimental conditions, results show that GPS allows the highest performance. However, accelerometer, pedometer and gyroscope supporting the baseline classifiers use a fraction of the battery of GPS [9]. In the same conditions, the BIBO classifier based on BLE signal is significantly better than the baseline and presents similar or lower battery consumption patterns. Concerning the classifiers, we notice that RF performs generally better than MLP, while maintaining resilience to noise > 20% of the labels, during training. This rate corresponds to an average of more than one entire segment per user with wrong labels.

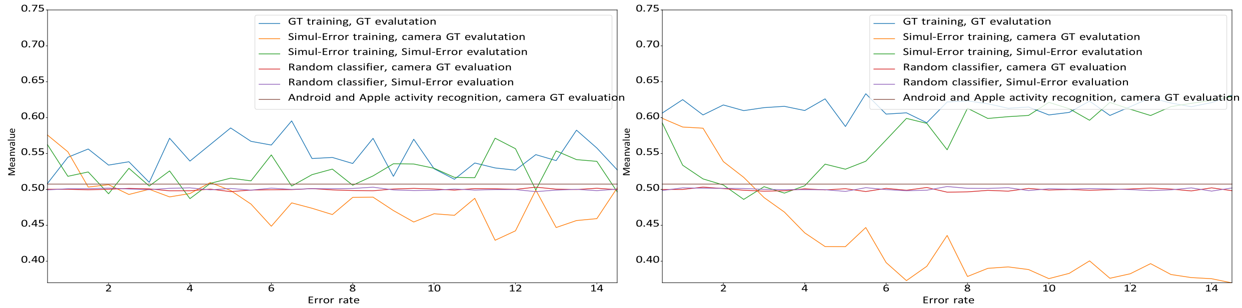

Figure 3: Random Forest, BIBO classification, AUC performance, left using BLE RSSI, right GPS (p-value with respect to random classifier on 0.5 value << 0.01)

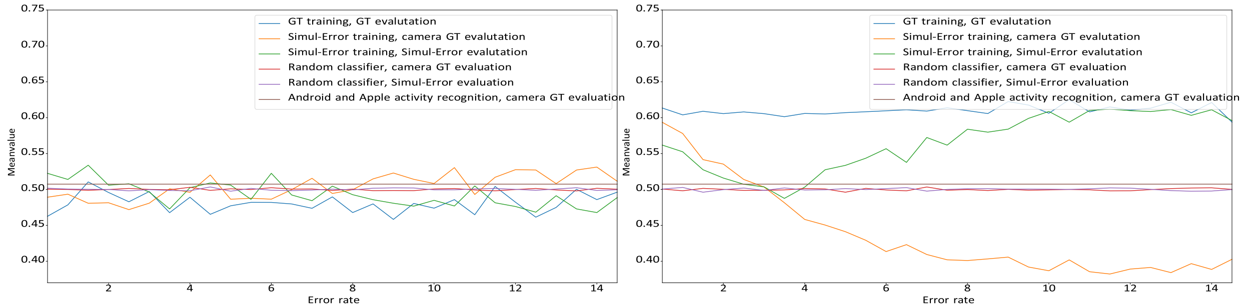

Figure 4: Multi Layer Perceptron, BIBO classification, AUC Performance, left using BLE RSSI, right using GPS (p-value with respect to random classifier on 0.5 value << 0.01)

Conclusion

This work presents a feasibility and sensitivity study about a BIBO system. This system is based on the smartphone onboard sensors including BLE RSSI signal detected from an external BLE beacons network. The system relies on a full-stack sensing platform deployed for data collection and constitutes the foundation for a context-aware smartphone-based travel survey. Empirical evidence shows that the exclusive use of BLE for BIBO classification is more effective than off-the-shelve classifiers based on inertial sensors such as accelerometer, gyroscope, and pedometer. GPS show higher potential but comes at the cost of higher battery consumption for the smartphone. Hence, future research should expose how to leverage available sensors to find the best tradeoff between BIBO accuracy and battery efficiency. As for now, BLE RSSI and GPS signals seem to enable independent BIBO classification, and reciprocal validation. Consequently, this platform could relieve passengers from BIBO ground truth collection and thus allow them to focus on other surveys, possibly BIBO-aware, and relevant for studying other aspects of their interaction with the autonomous vehicles.

References

[1] Corinna Li, P. Christopher Zegras, Fang Zhao, Zhengquan Qin, Ayesha Shahid, Moshe Ben-Akiva, Francisco Pereira, and Jinhua Zhao. Enabling Bus Transit Service Quality Co-Monitoring Through Smartphone-Based Platform. Transportation Research Record: Journal of the Transportation Research Board, 2649(1):42–51, 2017.

[2] Rico Krueger, Taha H. Rashidi, and John M. Rose. Preferences for shared autonomous vehicles. Transportation Research Part C: Emerging Technologies, 2016.

[3] Natasha Merat, Ruth Madigan, and Sina Nordhoff. Human Factors, User Requirements, and User Acceptance of Ride-Sharing in Automated Vehicles. International Transport Forum, 2017.

[4] Karen Lucas. Transport and social exclusion: Where are we now? Transport Policy, 20:105–113, 2012.

[5] Arto O. Salonen. Passenger's subjective traffic safety, in-vehicle security and emergency management in the driverless shuttle bus in Finland. Transport Policy, 2018.

[6] S. Technische Universiteit Delft. and H.J.P. Timmermans. European journal of transport and infrastructure research. European Journal of Transport and Infrastructure Research, 14(4):412–424, 2014.

[7] Jasper Dekkers and Piet Rietveld. Electronic ticketing in public transport: A field study in a rural area. Journal of Intelligent Transportation Systems, 11(2):69–78, 2007.

[8] Apple CORE Motion API, 2021.

[9] Valentino Servizi, Francisco C. Pereira, Marie K. Anderson, and Otto A. Nielsen. Mining user behaviour from smartphone data: a literature review. 2020.

[10] Wolfgang Narzt, Stefan Mayerhofer, Otto Weichselbaum, Stefan Haselbock, and Niklas Hofler. Be-in/be-out with bluetooth low energy: Implicit ticketing for public transportation systems. Ieee Conference on Intelligent Transportation Systems, Proceedings, Itsc, 2015-:7313345, 2015.

[11] Ali Yassin, Youssef Nasser, Mariette Awad, Ahmed Al-Dubai, Ran Liu, Chau Yuen, Ronald Raulefs, and Elias Aboutanios. Recent advances in indoor localization: A survey on theoretical approaches and applications. Ieee Communications Surveys and Tutorials, 19(2):7762095, 2017.

[12] Kotanen, M. Hännikäinen, H. Leppäkoski, and T. D. Hämäläinen. Experiments on local positioning with bluetooth. Proceedings Itcc 2003, International Conference on Information Technology: Computers and Communications, page 1197544, 2003.

[13] Fazli Subhan, Halabi Hasbullah, Azat Rozyyev, and Sheikh Tahir Bakhsh. Indoor positioning in bluetooth networks using fingerprinting and lateration approach. 2011 International Conference on Information Science and Applications, Icisa 2011, page 5772436, 2011.

[14] Dongyao Chen, Kyong-tak Cho, Sihui Han, Zhizhuo Jin, and Kang G Shin. Invisible Sensing of Vehicle Steering with Smartphones. Mobile Systems , Application and Services, pages 1–13, 2015.

[15] Fazli Subhan, Halabi Hasbullah, and Khalid Ashraf. Kalman filter-based hybrid indoor position estimation technique in bluetooth networks. International Journal of Navigation and Observation, 2013:570964, 2013.

[16] Héctor José Pérez Iglesias, Valentín Barral, and Carlos J. Escudero. Indoor person localization system through rssi bluetooth fingerprinting. 2012 19th International Conference on Systems, Signals and Image Processing, Iwssip 2012, page 6208163, 2012.

[17] Abdallah Moubayed and Abdallah Shami. Softwarization, virtualization, machine learning for intelligent effective v2x communications. Ieee Intelligent Transportation Systems Magazine, page 14, 2020.

[18] Apple BLE Detection API, 2021.

[19] Andreas Bjerre-Nielsen, Kelton Minor, Piotr Sapiezynski, Sune Lehmann, and David Dreyer Lassen. Inferring transportation mode from smartphone sensors: Evaluating the potential of wi-fi and bluetooth. Plos One, 15(7), 2020.

[20] David F. Nettleton, Albert Orriols-Puig, and Albert Fornells. A study of the effect of different types of noise on the precision of supervised learning techniques. Artificial Intelligence Review, 33(4):275–306, 2010.

[21] Ricardo Barandela and Eduardo Gasca. Decontamination of training samples for supervised pattern recognition methods. In Francesc J. Ferri, José M. Iñesta, Adnan Amin, and Pavel Pudil, editors, Advances in Pattern Recognition, pages 621–630, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.
[22] Carla E. Brodley and Mark A. Friedl. Identifying mislabeled training data. Journal of Artificial Intelligence Research, 11:131–167, 1999.
[23] André L.B. Miranda, Luís Paulo F. Garcia, André C.P.L.F. Carvalho, and Ana C. Lorena. Use of classification algorithms in noise detection and elimination. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5572:417–424, 2009.
[24] Christopher M. Bishop. Pattern Recognition and Machine Learning. 2006.
[25] GG Cantarero and RM Jarabo. The area under the roc curve. Medicina Clinica, 106(9):355–356, 1996.