A2B: Identifying movement patterns from large-scale Wi-Fi based location data.

S.C. van der Spek, E. Verbree, M. Bon, X.A. den Duijn, B. Dukai, S.J. Griffioen, Y. Kang and M. Vermeer

Delft University of Technology, Geomatics for the Built Environment,
Faculty of Architecture and the Built Environment,
Julianalaan 134, 2628 BL, Delft, The Netherlands

Extended Abstract

The distribution of people in buildings, the occupancy of lecture-, work- and study places and the accessibility of facilities are essential information at university campuses who have to cope with limited and even shrinking budgets and huge, rising real estate costs. Only little insight is gained in both occupancy and movement patterns with traditional counting techniques and user-based questionnaires. Management teams state that rooms and facilities are hardly used, though staff and students complain about overcrowded facilities and limited flexibility. Actual and accurate data on a 24/7 scale with high-granularity is missing.

In general Facility- and Asset Management lacks efficient methods for real-time, comprehensive and high-granularity information of location, capacity and use of tangible and intangible assets. Asset management could benefit from more detailed, more accurate and longitudinal data on assets, providing more insight into efficiency and effectiveness on different levels of scale through time.

Existing technologies could provide a platform delivering those required insights. Navigation- and communication technologies such as GNSS, Wi-Fi, Bluetooth, RFID can be used to 'locate' users, estimate intensities and reveal patterns of movement and patterns of use. For Asset management indoor localisation is essential.

Technology

Wi-Fi is a widespread communication technology used by electronic devices to connect to a Wireless Local Area Network (WLAN) base station or to connect ad-hoc directly between devices. Wi-Fi may be used to obtain

Published in "Proceedings of the 13th International Conference on Location-Based Services", edited by Georg Gartner and Haosheng Huang, LBS 2016, 14-16 November 2016, Vienna, Austria.

internet access, to exchange data, to access an intranet or for sending data to devices like a printer. Today a large range of electronic devices is capable of using Wi-Fi including computers, laptops, smart-phones, tablets, digital cameras, audio players, printers, (video-)game consoles and sensors. Wi-Fi networks are offered in companies, private homes, cities and (semi) public spaces and also at university campuses. Eduroam is a worldwide standard for University Campus WLAN networks.

Wi-Fi cannot only be used as a technology to transfer digital data wirelessly, but also as a tool for Facility and Asset management or as a platform for location-based services (LBS): Wi-Fi Access Points (AP's) can be used as sensors to collect information of connecting devices, delivering dashboards with temporal data on intensity of devices based on the number of unique devices detected and patterns of movement based on detections of the same device at different access points.

Wi-Fi 'user' data can be obtained in two ways: (a) by scanners or (b) by the network.

ad. (a) Wi-Fi scanners register connection attempts from devices within range. Every enabled Wi-Fi device is continuously searching for Wi-Fi access points and therefore broadcasting its unique media access control (MAC) address. No connection between device and scanner is made and no data is exchanged. Scanners can be other Wi-Fi devices as well as (modified) Access Points.

ad. (b) When using the network Access Points the real -established- connections are used. Either the connection start- and endtime is logged, or the system regularly scans for connected devices. At TU Delft a dump of devices connected to Eduroam Access Points is made every five minutes for the whole campus for network management purposes. Personal information such as MAC address and network ID are immediately encrypted (hashed). The research projects described in this abstract use this anonymised data from the Eduroam network for spatio-temporal analysis. Only staff and students from the university connecting to the eduroam Wi-Fi network are incorporated in this research.

In the Geomatics Synthesis Project Wi-Fi is used in a campus-wide experiment to monitor flows and occupation patterns at the TU Delft Campus. Students worked for two months on three parallel projects :

(1) extracting presence of people at specific places;

(2) unravelling patterns of movement within buildings and between buildings on the campus; and

(3) identifying activities and irregular use based on Wi-Fi data.

In all projects, the same dataset is used. All project also had to cover four cross-cutting topics: Privacy, Validity and Accuracy, Representativeness and reflection on the system of Access points (data collection).

Project 2: A2B : Identifying Movement patterns from large-scale Wi-Fi based location data.

A company Wi-Fi network works as a company-broad roaming service for Wi-Fi enabled devices. This enables to collect data of connection to this system and based on that aggregate to flows of devices through the system. This assists both Asset management by providing insight on actual movements between facilities and users by potentially offering tools to facilitating their movement behaviour. The second project specifically focusses on analysis of movement within and between facilities. In several corporate situations, Wi-Fi is already used for measuring movement, i.e. IKEA, Airports such as Copenhagen and Schiphol and the University campus of the National University of Singapore. Research has been carried out by Meneses and Moreire (2012) but also at TU Delft (Kalogianni et al. 2015).

The aim of this research is to use location data based on Wi-Fi logs to conduct a mobility analysis producing knowledge about the use and interaction on the Delft University Campus level.

Methodology

The TU Delft campus is equipped with over 1700 Wi-Fi Access Points (AP's) distributed over 30 buildings. The AP's mainly cover the indoor space. All students and staff have free access to this system. The system scans and stores all connections every 5 minutes for every AP. All data is encrypted. The data includes a timestamp, encrypted user information and AP name. The AP is located in a specific building at a specific place.

The analysis of movement focusses on two levels: between buildings and between building parts. The latter research has been carried out at the Faculty of Architecture as this was the only place where the exact location of access points was provided.

For the analysis of the data, a distinction is made between static devices, laptops and mobile devices like Smartphones. The hypothesis is that Smartphones are more representative for analysing movement behaviour than laptops and tablets. The latter are usually offline while moving through or between buildings and therefore not observed by the system.

The methodology for processing the data is based on the sequence and duration of the connection to specific AP's. The less AP's the more likely a device is static. The more AP's the more likely a device is mobile. Also being 'invisible' is taken into account by defining being 'away' for at least an hour is called 'world'. Connections to the same AP with short intervals are grouped in blocks. This way the status and sequence of every device can be defined as either world (outside), moving (between AP's) or present (within the range of AP).

The next step of the research was to derive movement patterns from the data. The movement was analysed on building-part level by grouping AP's to specific regions.

Research

The dataset comprised two months of data. Within this dataset this team discovered almost 45.000 different users and more than 85.000 different devices. Only around 24.000 of these devices were classified as mobile, the other devices were either static or had insufficient sessions connected to the Wi-Fi to deliver movement patterns.

The movement patterns were visualised in general and delivered a clear relation between the lecturing schedule and movement: main peaks at the 1st hour (08:45), during start and end of lunch (12:45 and 13:45) and at the end of the 8th hour (17:45); Small peaks can be observed at the intermediate breaks around 10:45 (after 2nd hour) and 15:45 (after 6th hour).

The former results were, of course, expected based on the academic schedule. More interesting are the observations of movement behaviour between facilities and the indoor movement patterns in the building of the Faculty of Architecture and the Built Environment. In general, the main movement is from most faculties in direction of the Aula for lunch, but even bigger is the flow to the library. In Architecture the pressure on the central street, especially from the East-Wing is visible.

Conclusion

On campus level, the Rhythm between buildings can be shown rather clearly based on the Wi-Fi data. Specifically, relations between Faculties can be illustrated. At building level, the same methodology can be introduced by using an indoor network graph based on the floorplan. This successfully illustrated movement in the building, however clearly issues with overlap between AP's can be observed.

Recommendation

On building level, the range of AP's needs to be investigated and translated into algorithms to identify the correct movement patterns. The performance of the tool is depending on building form. Architecture has a specific, ideal form with three main distinguishable parts. This will not always be the case.

For this research, especially the description of the outdoor space of TU Delft was missing. It is advised to add AP's in the public domain of the University Campus to make a distinction between people leaving the premises of TU

Delft and people using the outdoor facilities of TU Delft, like the centrally located Mekelpark area.

Acknowledgement

This project is part of the Geomatics Synthesis Project (GEO1101) 'Rhythm of the Campus' which was carried out in Spring 2016. The data is collected by ICT of TU Delft and made available for the project. In a later stage – if no obstructions are encountered – the data will be made open-source.

The course was initiated and coordinated by Stefan van der Spek, director Geomatics for the Built Environment and Edward Verbree, section GIS-technology. Edward Verbree, Martijn Meijers and Wilko Quak mentored the three project teams. Bart Valks and Iljoesja Berdowski represented Facility Management of TU Delft. Alexandra den Heijer and Ruud Binnenkamp (Management in the Built Environment) assisted to the overall scope of the project.

References

Agouris, P., A. Croitoru, A. Stefanidis (2008) FEATURE INTEGRATION FOR GEOSPATIAL INFORMATION A Review and Outlook, Chapter 18 in Digital Government Volume 17 of the series Integrated Series In Information Systems, pp 353-376

Agrawal, Rakesh and Ramakrishnan Srikant (1994). "Fast algorithms for mining association rules". In: Proc. 20th int. conf. very large data bases, VLDB. Vol. 1215, pp. 487–499. URL: https://www.it.uu.se/edu/ course/homepage/infoutv/ht08/vldb94_rj.pdf (visited on 04/09/2015).

Agrawal, Rakesh, Tomasz Imieli ´ nski, and Arun Swami (1993). "Mining association rules between sets of items in large databases". In: ACM SIGMOD Record. Vol. 22. ACM, pp. 207–216. (Visited on 04/09/2015).

Anbukkarasy, G. and N. Sairam (2013). "Interesting Metrics Based Adaptive Prediction Technique for Knowledge Discovery". In: International Journal of Engineering and Technology 5.3, pp. 2069–2076. (Visited on 05/16/2016).

Conference on Business Intelligence and Financial Engineering, 2009. BIFE '09. International Conference on Business Intelligence and Financial Engineering, 2009. BIFE '09. IEEE, pp. 460–463. DOI: 10.1109/ BIFE.2009.110.

Dodge, Somayeh, Robert Weibel, and Anna-Katharina Lautenschütz (2008). "Towards a taxonomy of movement patterns". In: Information visualization 7.3-4, pp. 240–252.

Graux, H. (2011). Open government data: reconciling PSI re-- use rights and privacy concerns. European Public Sector Information Platform Topic Report No. 2011 / 3.

Heijer den, Alexandra (2012). "Managing the University Campus: Exploring models for the future and supporting today's decisions". In:

Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment". In: Computing In Science & Engineering 9.3, pp. 90–95.

Kalogianni, E et al. (2015). “Passive WiFi Monitoring of the Rhythm of the Campus”. In: Proceedings of The 18th AGILE International Conference on Geographic Information Science. AGILE, pp. 1–4.

Kulk, S., B. van Loenen (2012), Brave New Open Data World? International Journal of Spatial Data Infrastructure Research, Vol.7, 196-206 (link)

Loenen, B. van, J. de Jong, J.A. Zevenbergen (2008), Locating mobile devices; balancing privacy and national security, NWO Research report. (link)

Mahmuda, A., S. Karagiorgou, D. Pfoser, · C. Wenk (2015) A comparison and evaluation of map construction algorithms using vehicle tracking data, Geoinformatica, 19:601–632, DOI 10.1007/s10707-014-0222-6

Mautz, R. (2012) Indoor Positioning Technologies. Chapter 8: "WLAN / Wi-Fi" ETH Zurich, Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry.

Mautz, Rainer (2012). “Indoor positioning technologies”. PhD thesis. Habilitationsschrift ETH Zürich, 2012.

Meneses, Filipe and Alberto Moreira (2012). “Large scale movement analysis from WiFi based location data”. In: Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on. IEEE, pp. 1–9.

Orellana, D., A.K. Bregt, A. Ligtenberg, M. Wachowicz, (2012) Exploring visitor movement patterns in natural recreational areas, in: Tourism Management, 33(3):672–682 DOI 10.1016/j.tourman.2011.07.010

Parliament and Council of European Union (1995). Directive (EC) 95/46/EC. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML.

Pei, Jian et al. (2004). “Mining sequential patterns by pattern-growth: The prefix span approach”. In: Knowledge and Data Engineering, IEEE Transactions on 16.11, pp. 1424–1440.

Petrenko, A., Sizo, A., Qian, W., Knowles, A. D., Tavassolian, A., Stanley, K. and Bell, S. (2014), Exploring Mobility Indoors: an Application of Sensor-based and GIS Systems. Transactions in GIS, 18: 351–369. doi:10.1111/tgis.12102

Radaelli, Laura et al. (2013). “Identifying Typical Movements among Indoor Objects–Concepts and Empirical Study”. In: Mobile Data Management (MDM), 2013 IEEE 14th International Conference on. Vol. 1. IEEE, pp. 197–206.

Scassa, T. (2010). Geographical information as ‘personal information’ Oxford University Commonwealth Law Journal, 10(2), 185-214.

Spek, SC van der (2008). “Mapping Pedestrian Movement: Using Tracking Technologies in Koblenz”. In: Lecture Notes in Geoinformation and Cartography, pp. 95–118.

Zhang, Yuejin et al. (2009). “A Survey of Interestingness Measures for Association Rules”. In: International

Zhao, Shao et al. (2014). “Discovering People’s Life Patterns from Anonymized WiFi Scanlists”. In: Ubiquitous Intelligence and Computing, 2014 IEEE 11th Intl Conf on and IEEE 11th Intl Conf on and Autonomic and Trusted Computing, and IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UTC-ATC-ScalCom). IEEE, pp. 276–283.