Presence Analytics: Discovering Meaningful Patterns about Human Presence Using WLAN Digital Imprints

Muawya H. Sarnoub Eldaw
Dept. of Computer Science
Birkbeck, University of London
eldaw@dcs.bbk.ac.uk

Mark Levene
Dept. of Computer Science
Birkbeck, University of London
mark@dcs.bbk.ac.uk

George Roussos
Dept. of Computer Science
Birkbeck, University of London
g.roussos@dcs.bbk.ac.uk

ABSTRACT

In this paper we illustrates how aggregated WLAN activity traces provide anonymous information that reveals invaluable insight into human presence within a university campus. We show how technologies supporting pervasive services, such as WLAN, which have the potential to generate vast amounts of detailed information, provide an invaluable opportunity to understand the presence and movement of people within such an environment. We demonstrate how these aggregated mobile network traces offer the opportunity for human presence analytics in several dimensions: social, spatial, temporal and semantic dimensions. These analytics have real potential to support human mobility studies such as the optimisation of space use strategies. The analytics presented in this paper are based on recent WLAN traces collected at Birkbeck College of University of London, one of the participants in the Eduroam network.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless communication; I.6.5 [**Model Development**]: Modeling methodologies

Keywords

Human Mobility, WLAN Traces, Eduroam, Presence Analytics

1. INTRODUCTION

The increasing advancement in wireless technologies together with the widespread use of new generations of faster and more powerful mobile devices has greatly improved the ability of people to access information while moving about in their daily lives. This increasing accessibility to digital information has the potential to generate vast amounts of detailed information, providing an invaluable opportunity to study different aspects of presence and movement behaviours of people within a given work or study environment. Furthermore, the increasing level of connectivity to information sources is affecting our environments and the way they operate and therefore, it is essential that we build real-time monitoring systems as well as theoretical frameworks to understand how people's presence and its dynamics reshape the structures of our environments. With such measurements put in place, we can discover hidden patterns of behaviour at both the collective and at the individual user levels, thus increase our understanding about people's presence, and in turn, improve our ability to make informed decisions when we plan for our environments.

In the context of this paper, we analyse the mobility traces generated by users accessing the wireless network at Birkbeck College, University of London. Birkbeck is a full participant of Eduroam (Education roaming), a WLAN service developed for the international education and research community, that gives secure, worldwide roaming access to the Internet [1]. The findings reported in this paper are the result of the analysis carried out on the Eduroam access traces, for the period from the first of October 2013 to the 10th of April 2015. In comparison to most data sets used in previous Eduroam based studies, this data set is larger in size with respect to its number of users as well as the number of days it spans [10, 2].

This paper investigates the human presence within an academic environment and examines four types of behavioural patterns to this end. These types of patterns correspond to the four different aspects of the data: the social, the spatial, the temporal and the semantic aspects. For each of these aspects, we define a list of metrics, which we utilise to interpret the observed behaviour captured in the data.

The paper makes the following two contributions:

1. It presents a comprehensive analysis about the human presence within a university campus. It investigates the four types of patterns contained in the data: the social, the spatial, the temporal and the semantic patterns, giving an insight into how people presence shapes the dynamic structure of such an environment. Although there are numerous previous works investigating the network usage of users in WLANs [7, 6, 3], there was no attempt to analyse the four aspects of human presence in one study. In the next section, we provide a succinct summary of some of these research efforts, which generally concentrate on characterising the network usage utilising one or two of the data aspects at most.
2. To our knowledge, the analysis provided in this paper is based on the most recent data set - compared to data sets used in previous Eduroam research - and thus reflects the current behavioural trends in WLAN usage in a university setting. With exception to the data set used in [4], the data set used in this research is much larger in size compared to

ICC '16, March 22-23, 2016, Cambridge, United Kingdom
© 2016 ACM. ISBN 978-1-4503-4063-2/16/03...$15.00
DOI: http://dx.doi.org/10.1145/2896387.2896438

data sets used in previous Eduroam based studies [10]. A further distinguishing property of the data set used in this work is the fact that a larger proportion of users come from other universities in the vicinity as opposed to being affiliated with Birkbeck College.

The rest of the paper is organised as follows: In Section 2, we review the related work. In Section 3, we describe the essence of presence analytics, giving the definitions of the concepts used in this work as well as describing the metrics used for discovering the different patterns of human presence contained in the data. In Section 4, we give a description of the data set we used for the analysis, followed by the evaluation of results. We provide a comprehensive discussion about the results in Section 5. Finally, in Section 6, we give our concluding remarks and a brief description of our future work.

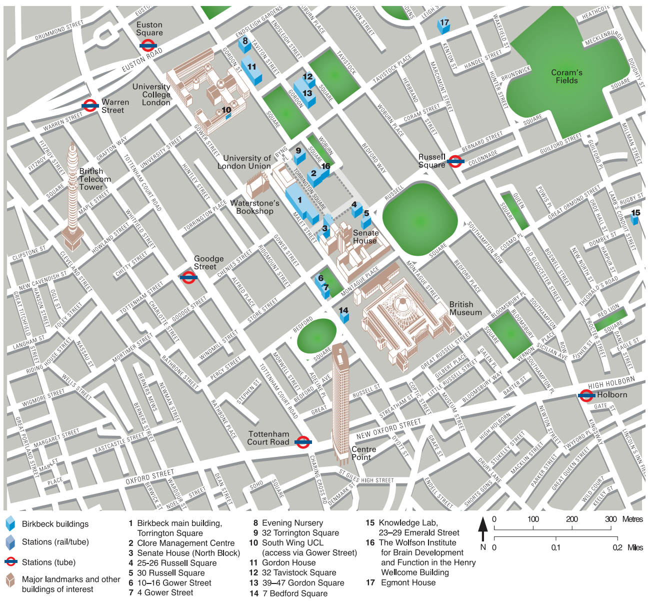

Figure 1: The Location of Birkbeck's Bloomsbury Campus in Central London.

2. RELATED WORK

There are numerous studies that investigated the possibility of using WLAN traces to get an up-to-date perspective of the human dynamics within an academic institution. We examine some of these studies with respect to the four data aspects discussed in the previous section, i.e. the social, the spatial, the temporal and the semantic aspects.

In a study discussed in [8] it was shown that it is possible to identify social groups amongst users. The study was based on WLAN mobility traces that were collected over a period of one month. In the same study, it was shown that male and female session duration can be significantly different. In [5], which discusses a study involving university students, Eagle and Pentland identified activity patterns related to the users daily behaviour. They further discovered that the daily patterns can be associated with the user's major of study and, in turn, linked to the level of employment. In [9] the authors estimated the long-term network usage among different access points. By being able to predict the future locations of users, and thus predict the distribution of future user locations, and in turn estimate the redistribution of loads among neighbouring Access Points in those locations.

In [12] the authors present a study in which a small, but carefully designed, WLAN was used to investigate the usage of the network before an expansion plan was drawn. The main goal of the study was to find out information about the time, the location and the level of usage when the users connected to the network. In this study the data set was collected in early 2003 over a period of only one week. In [11], which presents the finding of a relatively larger study as opposed to those aforementioned studies, the authors investigated the growth of the network from the spatial and temporal perspectives. The study discussed in this paper was carried out at the university of North Carolina and the data set used was collected in the period from October 2004 and April 2005.

With exception to [4], recent Eduroam studies provide analysis using small data sets of WLAN traffic traces. Although a large data set was utilised in [4], the analysis provided concentrates on the temporal aspect of the data and excludes the other dimensions of the human presence. In fact, all of the aforementioned studies in unison, do not provide a comprehensive analysis of the four dimensions of the human presence within an academic environment. The analysis provided in these studies overlooks one or more aspects of the human presence, in particular the semantic aspect, which has not been discussed in any of them. The analysis and the discussion presented in this paper is an attempt to bridge this gap. It is based on a large amount of WLAN traces, recently collected at Birkbeck College, which is one of the participant universities in Eduroam. Furthermore, this analysis provides an up-to-date view about the current trend in Eduroam usage.

3. THE ESSENCE OF PRESENCE ANALYTICS

3.1 Definitions

1. **Presence Analytics** is defined as the collection and the analysis of mobile data in order to find meaningful patterns about people's presence within a given environment.
2. **An Event** is defined as a group of one or more devices/users connecting to the network from a particular location within a given time interval.
3. **Revisit** is defined as the appearance of a user at a previously visited location or site.
4. **Duration of Stay** is the length of time that a user spend at a given location before moving to another location.
5. **Pattern of Event** is defined as a time series of occurrences of a given event, associated with a given time, e.g. evening or weekend pattern.

Within the context of this work, we rely on wireless network traces to gain information about human activity, in order to unravel the dynamic structure of the environment. Based on WLAN traces, activity patterns can be compared through time and space to reveal the dynamic structure of the environment. Presence Analytics allows for classifying the locations within a given WLAN environment, into functional clusters based on the time-line of human activity, providing valuable insights into the actual space use patterns within that environment. It provides new ways of looking at the structure of a given environment from a real-time perspective based on dynamic up-to-date records of human presence.

3.2 Data Sessionisation

Unless explicitly recorded, it is usual that WLAN access data does not include session duration information, i.e. the length of time an individual was using the network service to access information. Session duration information is essential for the type of analysis presented in this paper. It is used in computing the duration of stay, a core ingredient for quantifying the demand on space within the observed environment. It is important to highlight here that in this work, we prefer the term *duration of stay* over the term *session*, as it encompasses the social, the spatial and the temporal information required for presence analytics.

The data set we use in this work does not include the duration of stay information. Nonetheless, this data set has sufficient related information that can be utilised to compute approximate duration of stay.

When the user authenticates more than once within the same day, we compute their *approximate duration of stay* at that location using the sequence of authentications made by the user on that day. Practically, we apply a threshold based method utilising the length of the interval between the times of the user's authentications. This method is described as follows.

Using the timeline of authentications, we compute an ordered list of all the authentications made by the user. We also select a lower threshold value as well as an upper value - e.g. 1 and 110 minutes - and then apply the following procedure to compute the duration of stay:

1. We assume a duration of stay equals to the lower threshold value, i.e. one minute, for those users with a single authentication record per day.

2. If the user authenticated from the same location several times, and the difference between the times of two consecutive authentications is below the upper threshold value, then we assume that the user was present for the entire interval between those two consecutive authentications and thus the duration of stay is computed as the difference between the times of the two authentications.

3. If the user authenticated from the same location more than once and the difference between the times of two consecutive authentications is larger than the upper threshold value, then we assume that the user was present for a period equals to lower threshold value after the first authentication.

4. If the user moved to a different location, we assume that his duration of stay at the previous location is equal to the lower threshold value.

5. If the user has made a sequence authentications from the same location, and the difference between the times of two or more consecutive authentications is below the upper threshold value, we compute the accumulated sum of duration of stay for these consecutive authentications.

The lower and upper threshold values shown above, i.e. 1 and 110 minutes, adhere to the minimum and maximum session duration found in the literature [10].

3.3 Patterns of Human Presence

We distinguish between four types of patterns of presence in this work; each group corresponds to one of the aspects of the data discussed earlier in the Introduction section.

3.3.1 Social Patterns

The patterns discussed in this category represent the social perspective of the human presence. To examine such patterns, we utilise a collection of metrics which measure the influence of the social behaviour in the data. Some of these metrics capture where the users come from. For example, in this work we utilise some of the metrics to investigate the distributions of the users' study or work affiliation to see what social patterns can be extracted. The metrics address the following questions:

1. What are the institutions that the users are affiliated with.

2. Which affiliations take the top ranks in terms of the number of users.

3. How does the number of Birkbeck users compare to the numbers of users from other institutions particularly during the evening.

3.3.2 Spatial Patterns

The patterns in this category capture the spatial view of the human presence. The metrics used here analyse the users' behaviour from a spatial perspective. For example, we study, through these metrics, how the space is used - how the different locations are used by the users. We look at the impact of the division of the space into multiple sites. We examine the patterns of revisits to individual sites as well as the individual locations within the sites. Furthermore, we investigate a number of questions relating to the use of space at Birkbeck. The metrics address the following questions:

1. Which locations tend to be used the most by the top ranked affiliations

2. What times these locations are being used.

3. What locations do Birkbeck users use the most.

3.3.3 Temporal Patterns

In this category the metrics look at the data as time series and examine it for the existence of trends and seasonal patterns. We also examine the time that the user spend at a given location.

3.3.4 Semantic Patterns

The semantic dimension is an integral part of the framework that this paper is presenting in order to understand the influence of the presence of humans within a given environment. This framework is based on four main questions involving the four dimensions of the human presence. We describe the social dimension as the *who* dimension, the spatial dimension as *where* dimension, the temporal dimension as *when* dimension, and finally the semantic dimension can be described as the *why* dimension.

In this type of patterns we investigate the use of external information in giving meaning to the user's presence - in other words, we are trying to answer the question of *why* the user's behaviour was seen within the given social, spatial and temporal context. For example, why a student is seen in a specific lecture room at a given time or why a group of students and staff were seen at the coffee-shop between 12:00 and 1:00pm.

Further discussion on the semantic dimension will be given in Section 4.5.

4. DATA ANALYSIS

4.1 Birkbeck Data Set

Birkbeck College - a member college of University of London and a major provider of evening higher education. Based on the most recent available statistics, there are 16,463 students attending Birkbeck, with 38% studying for undergraduate qualifications. Most of Birkbeck students are part-time, with 88% of them enrolled on part-time programmes.

Birkbeck's Bloomsbury Campus in central London (See Figure 1), is located very close to campuses of other colleges of University of London, such as UCL and SOAS. This proximity to these other campuses was naturally translated in a large amount of collaboration between these universities. Today, Birkbeck's Bloomsbury campus is shared by thousands of academic, researchers and students from these universities on a daily basis. Birkbeck, is also one of the participant of Eduroam, a WLAN service developed for the international education and research community that gives secure, world-wide roaming access to the Internet [1].

The analysis presented in this paper is based on recent WLAN traces collected at Birkbeck. The data set we used here is a snapshot of the College's Eduroam access data for the period, from the 1st of October 2013 to 10th of April 2015. It contains 223 locations and 167272 users, who come from 2462 institutions and departments. The 223 locations given in this data set are divided between 11 of the 17 sites of the Bloomsbury campus (See Figure 1).

The data is divided into four categories: *authentication details*, *pre-proxy details*, *post-proxy details* and *reply datails*. User-ID, access location, timestamp and affiliation are the basic information for each processed record. Based on these records, we build new types of data representing the four aspects of the human presence: social, spatial, temporal and semantic aspects. The analytics presented in this work are based on this new data.

4.1.1 Data Privacy

All sensitive data items in this data set, such as the user's email/name and their device's MAC address, have been anonymised to allow the type of analytics provided in this work to be carried out without compromising the user's privacy. Furthermore, we do not attempt to use location in a way that compromises privacy e.g. via fingerprinting or display actual locations on maps. Although it is possible to do this but following ethical research guidelines we chose not to do so.

4.2 Evaluation of Social Patterns

Since the users in our data set are socially grouped by affiliation, we wanted to find out how they are distributed across these affiliations. As shown in Figure 2, which provides four different distributions: daytime, evening, weekdays and weekend, the distribution across the various affiliations is approximately a power-law. This "rich gets richer" trend means that most users belong to a small number of affiliations while the many more affiliations have a relatively small number of users.

4.3 Evaluation of Spatial Patterns

The 223 locations given in our data set are divided between 11 sites within the Bloomsbury campus. Since we are interested in finding out how the space is used, a lot of our efforts were focused on investigating the number of people visiting these sites and the locations within them. Because we are interested in the regular patterns of space use, we considered the revisits to the sites and excluded those single visits, i.e. visits by individuals who came to college only once, which add unnecessary noise in this case. As one can see from Figure 3, the users' revisits are exponentially distributed across locations. This means that the further we move across the locations towards the tail of the distribution, the faster the decrease in the number of revisits made to the location. At the thin tail of the distribution the locations are rarely visited.

4.4 Evaluation of Temporal Patterns

The main goal of this section is to analyse people's presence from a temporal perspective. We are interested in discovering the trends of regular presence as opposed to the occasional behaviours due to special events. Therefore, visits by individuals who only came to Birkbeck once are seen as special events, and thus excluded from this analysis to avoid the introduction of undesired noise. We use data decomposition to estimate the seasonal influence as well as any random noise. By removing these estimated components from data, we can reveal the temporal trend in people's behaviour.

Since the seasonal variation in the number of revisits is relatively constant over the period of time covered by the data, we consider an additive model for the decomposition:

$$Presence\ Data = Trend + Seasonal + Noise. \qquad (1)$$

This relatively constant seasonal variation is clearly visible in the bottom plots of the Figures 4 and 5 shown below.

4.4.1 Term-based Signature

We are interested in the tend, which illustrates the temporal variation in the number of revisits across the different academic terms. To extract such a trend, we divided the data into 13 weeks periods and applied an additive model to estimate the constituent behaviours such as the seasonality. Here, regular holidays such as Christmas and Easter holidays are considered seasonal events, which are captured well by the additive model. The plots given at the lower parts of the Figures 4 and 5 show these extracted seasonal events. The dipping points in the seasonality graph can be linked to such events.

The extracted termly tends for both Malet Street and Gordon Square sites show very similar patterns. In Figure 5, which shows the time series analysis for Gordon Square site, we see that the estimated trend shows a decrease from about 3000 revisits in the second period to about 500 revisits in the fourth period. This decrease preceded a steady increase to about 4000 revisits in the sixth period. In Figure 4, which shows the time series analysis for Malet Street site, we see almost an identical pattern to the trend shown at Gordon Square site. We see that the computed trend shows a decrease from about 5500 revisits in the second period to about 1000 revisits in the fourth period, followed by a steady climb to about 7500 revisits in the sixth period. The computation and the time series analysis carried out to produce the plots shown in the Figures 4 and 5 was performed in R.

In Figure 6, we see that the distribution of the time a revisiting user spend, on average, at a given location is approximately a power-law. This means that there is only a small number of high duration of stay while the many more remaining durations are for a relatively very short times.

4.5 Evaluation of Semantic Patterns

In Section (3.3), we focused our discussion on the core three dimensions of the human presence; the social, the spatial and the temporal dimensions. This was followed by a brief note about the semantic dimension and how it can potentially add meaning to the patterns linked to those three core dimensions. In this section, we

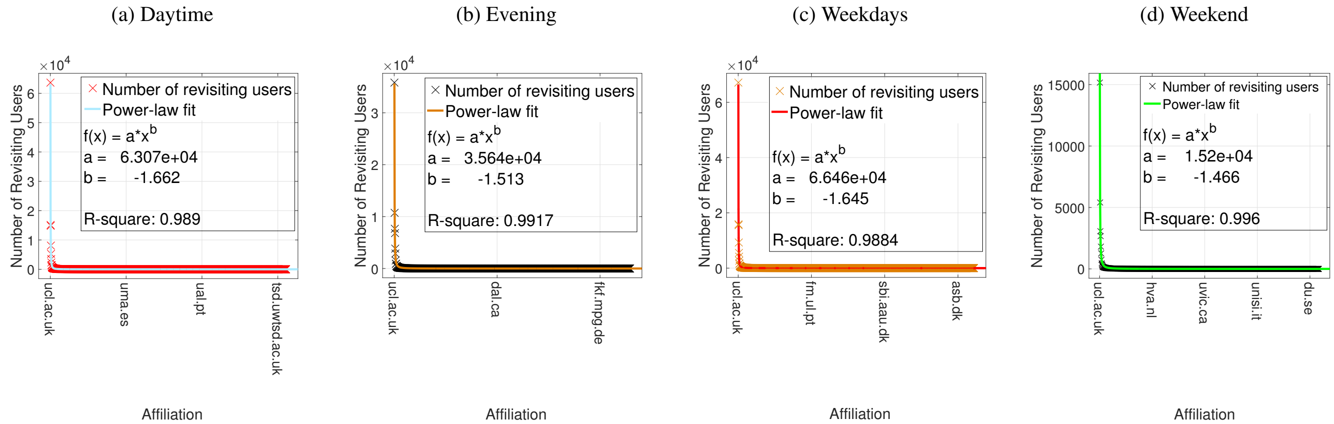

Figure 2: Distributions of number of revisiting users by affiliation. Each plot presented in this figure show the fitted distribution across all visited affiliations for the chosen period, e.g. Daytime. Revisiting users are those who made more than one visit to Birkbeck College

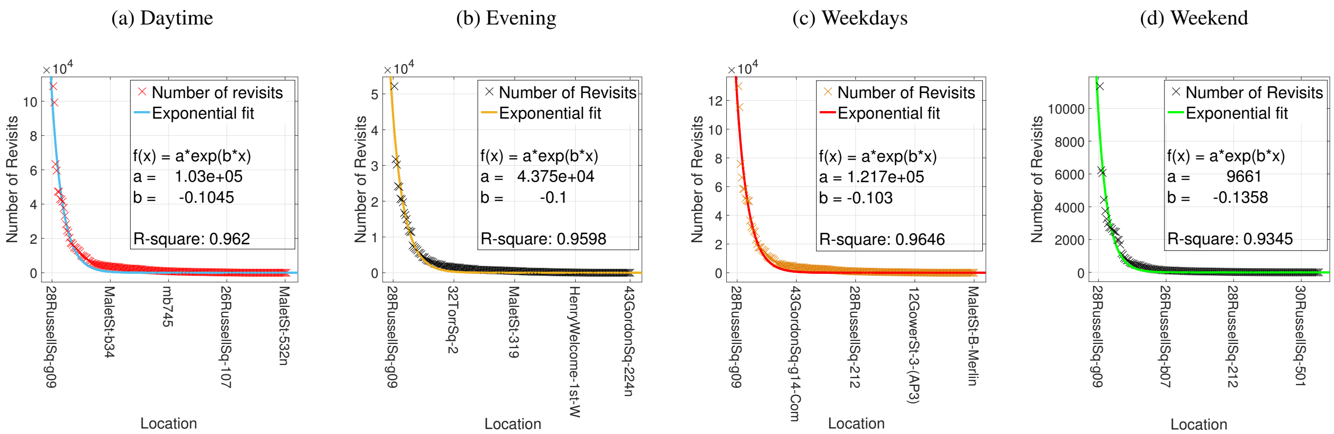

Figure 3: Distributions of number of revisits by location. The individual plots presented in this figure show the fitted distribution across all locations visited during the chosen pattern period, e.g. Daytime. The number of revisits to a given location is computed as the number of visits decreased by one.

study the semantic influence using an example in which external data from the teaching timetable is utilised. In this example, we use a two-step process, which can be described as follows:

1. We obtain the information, about the pattern we would like to discover, from the external source.
2. We analyse the data for temporal patterns that possess distinguishing characteristics which match the information obtained from the external source.

Of course, this process can be completely reversed, where we begin with the extraction of the patterns from the data and then map or link those discovered patterns to the information obtained from the external source in order to justify the existence of these patterns.

4.5.1 *Linking classrooms to the subjects taught through the teaching timetable:*

By combining social, spatial and temporal information with external data, such as the teaching timetable, we can identify groups of users whose presence at a given location is based primarily on the semantic influence as opposed to any other reason. To demonstrate this, we conducted an experiment in which we analysed the socio-spatio-temporal patterns for one of the computer labs at Malet Street site. From the teaching timetable, we selected the XML module, which was taught by the first author of this paper. This module had regular teaching sessions that ran from 18:00 to 21:00 every Monday in the period from the 12th of Jan to the 9th of March 2015. Socially, there was a total number of 14 individuals who attended the college for this module and had recorded traces within the data set. Note here that in order to map these patterns onto the teaching timetable we only consider the extracted patterns for the period of time that the teaching sessions cover. The result of this experiment is given in Figure 7, which shows the attendance of individuals from the selected group. In this experiment, the aggregate number of those who actually attended the sessions was 13. There was only one individual, who had no traces within the extracted patterns for the observed spatiotemporal context in which the sessions were taking place. Interestingly, this individual had traces in other locations within the temporal context of these sessions. The failure to detect the attendance of this individual can be attributed to the mobile device of this user being switched off while the session was taking place. The example given above is a proof of concept for the influence of the semantic dimension and the value that it adds to the other dimensions of the human presence analytics. Of course this paper

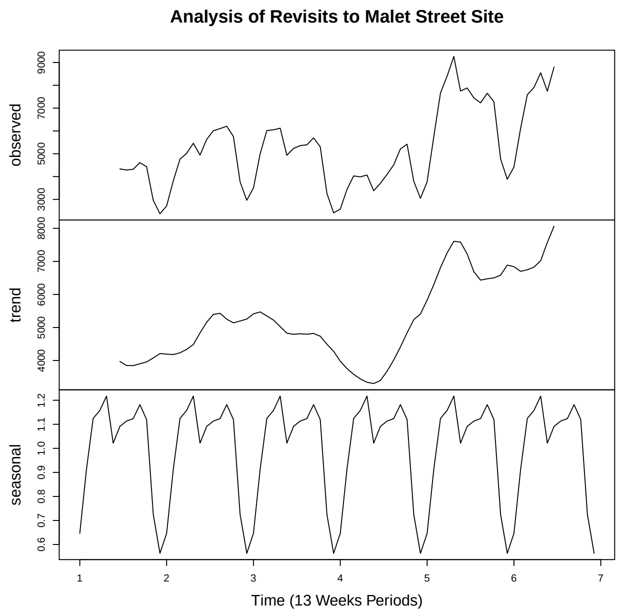

Figure 4: Time series analysis of number of revisits to Malet Street site. In this figure, the top plot shows the original time series in which the data is divided into 13 week periods, the plot second from top shows the estimated trend, and the bottom plot shows the estimated seasonal constituent

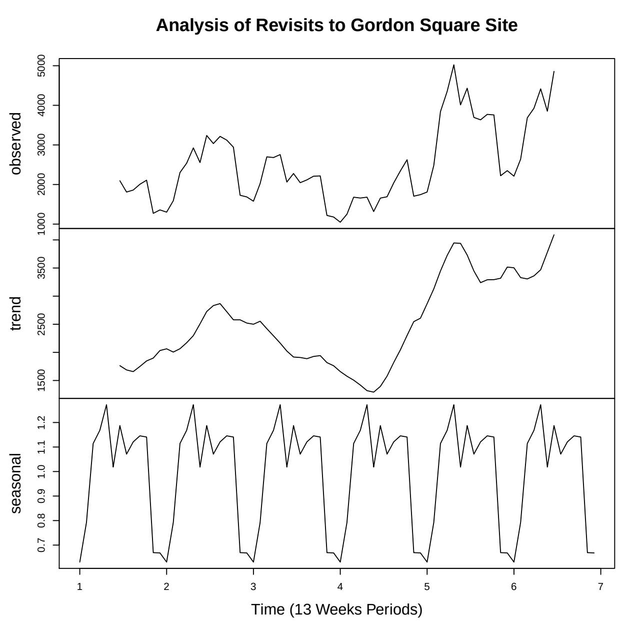

Figure 5: Time series analysis of number of revisits to Gordon Square site. In this figure, the top plot shows the original time series in which the data is divided into 13 week periods, the plot second from top shows the estimated trend, and the bottom plot shows the estimated seasonal constituent

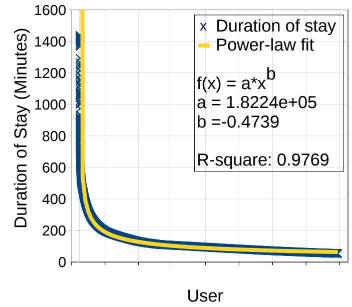

Figure 6: Distribution of duration of stay.

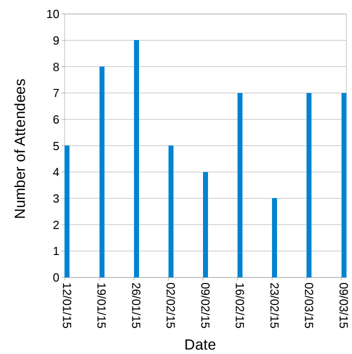

Figure 7: Attendance of the XML Module sessions as seen through traces of WLAN activity.

is mainly about the core three dimensions but no doubt the semantic dimension is an important perspective that provides invaluable insight about people's presence behaviour, and thus had to discuss, at least briefly, in this paper.

5. DISCUSSION OF RESULTS

Although this paper does not include how the users utilise the wireless network, in terms of the applications being used, the type of information being transferred and the rate in which the transfer happens, the analysis it gives already provides excellent insights about the way people interact with their environment. This suggests that reports showing accurate information about the presence and movement of people within the environment can be a very useful tool for planners when making decisions about the restructuring of the environment and how it operates. As people's connectivity to information sources becomes more ubiquitous and widespread, utilising such tool will become more common, perhaps as an additional tool to the more traditional ones such the expensive surveys and the classical static maps and drawings.

5.1 Limitation of Eduroam

Eduroam has the advantage that it is pervasive throughout the university and requires a single setup for authentication. Similar to most WLAN services, without registering the mobile devices with the service it is not possible to obtain any activity traces that can be linked to the users of those devices. In the experiment about the attendance of the XML sessions (discussed in 4.5), a larger proportion of those who attended the class did not have traces in our data set. These individuals might have not registered with the Eduroam service and thus we could not track their activity and determine their whereabouts when the teaching sessions were taking place. This shows a limitation in utilising Eduroam for mobile devices tracking within a given environment.

6. CONCLUSION AND FUTURE WORK

6.1 Conclusion

We provide a comprehensive analysis, about the human presence within a university campus. We investigate the four types of patterns contained in the data: the social, the spatial, the temporal and the semantic patterns, giving an insight into how people presence shapes the dynamic structure of such an environment. Our analysis is based on WLAN activity traces collected at the Birkbeck College. These traces are the most recent Eduroam data in comparison to data used in other previous Eduroam research, and thus the provided analytics reflect the current behavioural trends in WLAN usage in a university setting.

From a social perspective, our analysis reveals that the distribution of revisiting users across the various affiliations is approximately a power-law. The various patterns investigated: daytime, evening, weekdays and weekend, show that most users belong to a small number of affiliations while the many more affiliations have a relatively small number of users. From a spatial perspective however, we discovered that the users' revisits are exponentially distributed across locations. From a temporal perspective, the extracted termly tends show very similar patterns of revisit. The trend generally seems to gradually decrease reaching its lowest point in August, followed by a steady climb that reaches a peak at the end of November or the beginning of December.

To demonstrate the influence of the semantic dimension we show how combining social, spatial and temporal information with external data can give meaning to the user's behaviour. As a proof of concept, we give an example in which we utilise the teaching timetable to interpret the presence of a group of students attending a three hour regular session that took place on a weekly basis in one of the computer labs at Malet Street site.

We also discuss how Presence Analytics provide excellent insights about the way people interact with their environment, and how these analytics can be a very useful tool for planners when making decisions about the restructuring of the environment and how it operates. Furthermore, we discuss the limitation in utilising WLAN services such a Eduroam, for mobile devices tracking within a given environment.

6.2 Future Work

6.2.1 *Predicting future presence*

Motivated by the finding that there is high time regularity in human presence (See seasonality in Figures 4, 5), we assume that users have preferences with respect to the locations they visit. We have presence behaviour that can be described using an additive model. Furthermore, this behaviour can be adjusted by removing the seasonal influence. This means we may apply exponential smoothing to predict future behaviour, at least for the short term [13].

6.2.2 *Granular Social Patterns*

Social relationships is an integral part of every community and no doubt that numerous relationships and social networks exist between the people of the same university community. The people at Birkbeck are no exception to this. Unfortunately, there is no explicit information, about social relationships other than the affiliation, directly available from our data set. However, through the day-to-day social activities such as lectures, seminars and regular meetings, we have strong evidence about the existence of finer-grained relationships as opposed to the user affiliation. Extracting such relationships and measuring their influence on the other dimensions of the human presence is a key investigation of our future work.

7. ACKNOWLEDGEMENTS

We thank Kevin Brunt from Birkbeck IT Services who extracted the data set for us for the purpose of this research.

8. REFERENCES

[1] eduroam at Birkbeck - IT Services. Birkbeck, University of London. http://www.bbk.ac.uk/its/services/wam/eduroam. Accessed: 2015-09-14.

[2] A. Allahdadi, R. Morla, A. Aguiar, and J. S. Cardoso. Predicting short 802.11 sessions from radius usage data. In *Local Computer Networks Workshops (LCN Workshops), 2013 IEEE 38th Conference on*, pages 1–8. IEEE, 2013.

[3] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan. Characterizing user behavior and network performance in a public wireless lan. In *ACM SIGMETRICS Performance Evaluation Review*, volume 30, pages 195–205. ACM, 2002.

[4] N. Cruz, H. Miranda, and P. Ribeiro. The evolution of user mobility on the eduroam network. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*, pages 249–253. IEEE, 2014.

[5] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and ubiquitous computing*, 10(4):255–268, 2006.

[6] T. Henderson, D. Kotz, and I. Abyzov. The changing usage of a mature campus-wide wireless network. *Computer Networks*, 52(14):2690–2712, 2008.

[7] D. Kotz and K. Essien. Analysis of a campus-wide wireless network. *Wireless Networks*, 11(1-2):115–133, 2005.

[8] U. Kumar, N. Yadav, and A. Helmy. Gender-based feature analysis in campus-wide wlans. *ACM SIGMOBILE Mobile Computing and Communications Review*, 12(1):40–42, 2008.

[9] J.-K. Lee and J. C. Hou. Modeling steady-state and transient behaviors of user mobility: formulation, analysis, and application. In *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing*, pages 85–96. ACM, 2006.

[10] M. M. Mulhanga, S. R. Lima, and P. Carvalho. Characterising university wlans within eduroam context. In *Smart Spaces and Next Generation Wired/Wireless Networking*, pages 382–394. Springer, 2011.

[11] M. Papadopouli, M. Moudatsos, and M. Karaliopoulos. Modeling roaming in large-scalewireless networks using real measurements. In *Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks*, pages 536–541. IEEE Computer Society, 2006.

[12] D. Schwab and R. Bunt. Characterising the use of a campus wireless network. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 862–870. IEEE, 2004.

[13] P. R. Winters. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3):324–342, 1960.