Human Behaviour Detection Using GSM Location Patterns and Bluetooth Proximity Data

Muhammad Awais Azam, Laurissa Tokarchuk, Muhammad Adeel
Department of Computer Science and Electronic Engineering,
Queen Mary University of London,
UK, E1 4NS
{muhammad.azam, laurissa.tokarchuk, muhammad.adeel}@elec.qmul.ac.uk

Abstract— **Human behaviours are multifarious in nature and it is a challenging task to predict and learn from daily life activities. The profusion of Bluetooth enabled devices used in daily life has created new ways to analyze and model the behaviour of individuals. Bluetooth integrated into mobile handsets can be used as an efficient short range sensor. The aim of this research work is the detection of unusual human behaviours from cell tower and Bluetooth proximity data using neural networks. The primary purpose is to find anomalies in individual's daily life routines that will further help us to detect and predict unusual behaviour of elderly people and patients such as dementia patients.**

Keywords- Behaviour; Cell tower ID; Bluetooth proximity; Neural Network; Jaccard Index

I. INTRODUCTION

Detection and prediction of human behaviour is a hot issue nowadays in social research circles. Modelling human behaviour such as individual routines from proximity data and relations with gathered data of daily life activity patterns is an emerging realm in ubiquitous computing. There can be different sensing devices e.g., Radio Frequency Identification (RFID), motion sensors, GPS enabled tracking devices [5], and other context aware devices that can be used for real time proximity detection and daily life data gathering purposes. In particular, devices such as mobile phones provide a rich platform for various forms of data gathering by using its integrated sensors such as Bluetooth ID, digital camera, microphones and GPS transceivers. These sensors can give an individual's location, movement and proximity information for the whole period of cell phone usage. Specifically, Bluetooth radios are frequently incorporated into mobile devices [3].

Human behaviours and activities are analyzed by researchers using different sensor devices such as accelerometers, digital cameras and microphones. Frameworks have been presented to identify close proximity social behaviours [16], group actions in meetings [17] and audio visual perception of a lecture in smart environment [18]. In most studies, the majority of the sensing devices that are used in limited environment restrict their usage and thus it is only useful when activities take place in their proximity. This does not suit our case as we are interested not only indoor and limited environments but also outdoor movement and activities.

The enormous penetration of Bluetooth devices have enabled them to be used as a personal identifier. Many researchers have exploited this capability by using the mobile as a sensing device. The mobile phone nowadays is an indispensable part of our society with many integrated sensors. Researchers have investigated using these sensors in social proximity sensing [8][19], social behavioural modelling and routine classification [1][2][3][4] and movement prediction [6][7]. The significance of these studies is that they have identified new techniques to recognize an individual's behavioural patterns and abnormal movements. In [1][2], Author Topic Model (ATM) and hierarchical Bayesian topic models like Latent Dirichlet Analysis (LDA) are used for routine classification. A framework for daily life activity recognition based on the user's location and group affiliation is then presented. In [6][7], neural networks are used to detect and predict user movement based only on cell tower IDs. Our work is similar in one aspect with their work and that is; we have also utilized the probabilities of user being in different locations. Difference between our work and the work presented in [6][7] is that we have used real time data for our experiments and used both cell tower ID and Bluetooth proximity data.

This work is an extension of [9], in which repeated patterns and behaviour of an individual was detected by using n-gram technique and considering only Bluetooth proximity data. The primary purpose of this research is to detect unusual daily life activity patterns and individual behaviours that deviate from their normal routines in order to aid in the detection of abnormal behaviour such as wandering behaviour, a behavioural disorder in dementia patients. Another aim is to determine the reliability of behaviour detection using only Bluetooth proximity data. So, behaviour detection is done by considering the record of cell tower ID's and Bluetooth proximities because of the easy and economical availability of

 ISBN: 978-1-61208-100-7

Bluetooth enabled devices such as mobile phone as sensing device. This detection of Bluetooth Proximate devices shows the regularity of user's behaviour as discussed in [4]. According to [4], if the user in his daily life, repeat the activities and routines with less change, it will be known as 'low entropy' behaviour. While a more change in daily routine patterns is considered as 'high entropy' behaviour. Now, if we consider the elderly people and patients specifically dementia patients, they have somewhat fixed and regular routines to follow [20] that make those individuals a low entropic user based on [4]. Therefore our interest is to study primarily the users with low entropy from the reality mining dataset [4].

The results presented here use the reality mining dataset collected at MIT for the year 2004-2005. Nokia 6600 cell phones were used to record the data of 100 users over duration of 9 months. This research uses cell tower ID and Bluetooth proximity data to analyze the routines and behaviour of an individual that deviate from their normal routines. This paper presents the techniques used and corresponding analyses of the data that show the level of behavioural abnormality of individual's routine by using cell tower IDs and Bluetooth proximity information.

The rest of the paper is as follows: Section-II contains related work on abnormal activity detection and usage of Bluetooth as a sensing device. Section-III explains the methodology of our analysis that we have adopted to get the results. Section-IV discusses the results and Section-V contains the summary of the work and notes on the direction planned for our future work.

II. State of the Art

Detection of abnormality in human behaviour is very intricate and challenging task for researchers and has been in the past. Recently, with improvements in network systems and information technology, people have more easily been able to study the behaviours and activities of humans. Researchers have tried to detect the abnormal routines and daily life patterns of an individual inside the home and restricted environments [10][13]. Majority of work in this area has used sensing devices that either have short range of detection, less battery power and storage, or not very common that every person can use it without adding extra hardware, which is not possible for the scenarios outside the home. In [10], researchers have presented a framework for the detection of unusual human behaviour inside an intelligent house that is different from our case as we are considering the scenarios outside the home as well. They used motion sensors to detect the activities and unusual patterns based on Markov Chain. Vector quantization is used to reduce the sensor states and the change between these states is observed by transition probability. They detect the unusual behaviour by computing the distance between the state transition probabilities or by the likelihood of user action. The distance between the state transition probabilities was calculated by using either Kullback-Leiber distance or Euclid distance.

In [13], researchers detect abnormal event in solitary elder's daily life by mining the related data gained by sensors. They employ the association rules finding algorithm with time cluster to analyze the elder's activities. In first step, they cluster each item of elder activity with time and then in the second step, all frequent item sets were found and strong association rules were created. Researchers in [15] work on the recognition of abnormal activities based on the Hierarchical Dirichlet Process Hidden Markov Model (HDP-HMM). They incorporate a Fisher Kernel into the One-Class Support Vector Machine (OCSVM) to filter out the most likely normal activities. Then from those normal activities, they derive a model to detect abnormal activities and tried to reduce false positives. In [14], a model for abnormal behaviour detection is presented. That model considers user's location based on the cell tower ID and used Dynamic Bayesian Networks (DBN) to predict user's location. They proposed an X-Factor model, which is a DBN with a hidden variable. User's location according to this model not only depends on the hour of the day and day of the week but also this latent variable that represents the abnormal behaviour. In our work, we have applied neural network to detect abnormal behaviour of an individual by using cell tower ID and Bluetooth proximity data. This detection of behaviour will help us to aid elderly people and dementia patients.

III. Methodology

The aim of this research is the detection of abnormal behaviour in an individual's daily routines in order to aid in the detection of unusual behaviours in patients such as dementia patients, by using cell tower ID data and Bluetooth proximity data. This section starts by evaluating the use of only cell tower ID data and describes the methodology used in this research. Same methodology is then applied on the Bluetooth proximity data with some changes that are discussed in more detail in results section.

Cell tower ID gives information about the user's location and movement. The cell tower ID data that is used in this study is classified into four different locations; i.e., Home (H), Work (W), Elsewhere (E) and NoSignal (N). This data is divided into twenty four time slots. Each time slot is represented by the associated presence information of the user (H, W, E, and N) during the one hour period. The presence of user at specific location depends on the hour of the day and day of the week. For example, if the user has a regular routine of going to office, then location of the user at 10a.m on Saturday morning can not be the same at 10a.m on Monday morning. The daily life activities of an individual depend on the entropy level of the user as discussed in [4]. If the user is a low entropy user, his routines do not change much as compared to high entropy users, whose routines and activity

 ISBN: 978-1-61208-100-7

patterns change continuously. Figure 1 show the basic architecture used to get the behaviour of an individual. Data Base (DB) contains the classified information of cell tower ID data into H, W, E and N.

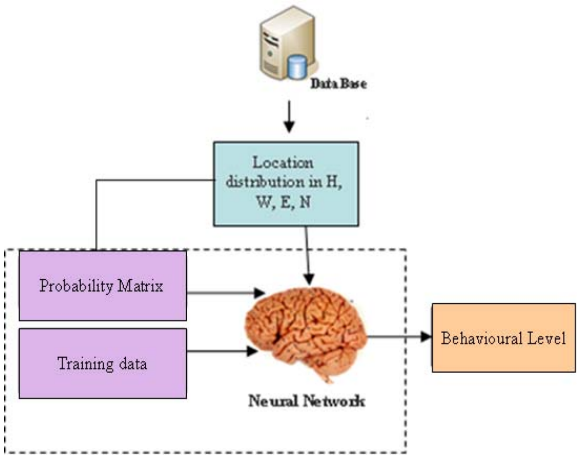

Figure 1. Data Processing Design

A probability matrix is generated depending on the hour of the day and day of the week from this classified information. This means that every entry of this matrix depends upon the specific hour of the day and whether it is a weekday or a weekend. This matrix is utilised for preparing training data for the neural network that is used for the detection of level of abnormality in the user behaviour. The neural network used for this purpose is Multilayer Perceptrons (MLP). There are four inputs and one output of this MLP. Output is the behaviour of the user and the format of the input used for this neural network is:

$$[Loc_i, Hour_j, Day_k, Abn\_Level_m]$$

where $Loc_{i,}$ gives the location i.e., H, W, E, N. $Hour_j$, gives the hour of the day i.e., between 1 and 24. Day_k, gives the day of the week, i.e., between 1 and 7. Abn_Level_m, gives the behavioural levels. Four different levels are assigned to behaviour depending upon the probability of being at any of the four places on a specific day and hour, shown in Table 1.

This generates twenty four samples of training data for one day. So for each user, total training samples are (24 x number of days). 70% of this training data is used for training the neural network whilst the remaining is used for cross validation and testing purposes. Training of the neural network is done till the cross validation error becomes less than 0.02, by using Mini-Batch training process [11], an advantage of using Mini-Batch training is that it is a compromise between batch and incremental training. Output of this neural network will give the level of abnormality of an individual for each hour of the day.

TABLE 1

Probability	Behaviour
$0 < p < 0.25$	Abnormal
$0.25 < p < 0.5$	Low Abnormal
$0.5 < p < 0.75$	Average Normal
$0.75 < p < 1$	Normal

IV. RESULTS

The results discussed here are of one user with the entropy level 23.06, calculated by using the Shannon's entropy equation given below.

$$H(x) = -\sum_{i=1}^{n} p(i) \log_2 p(i)$$

One month data is used to detect the behavioural levels of the user after training the neural network on about 70% of the data available for this specific user. First, only the cell tower ID data is used to detect the behaviour of the user. Figure 2 shows the daily distributions of (H, W, E and N) transitions based on cell tower ID data of one month that is further used to detect the behaviour of the user through neural networks.

Figure 3 shows the comparison of behaviour of an individual for two days. The trained neural network provides the behavioural levels for twenty four hours. As the entropy level of the user is quite low, this figure shows that most of the time the behaviour of the user is average normal. Now look at day-10 in Figure 2, there is an unusual detection of 'Elsewhere' during 5-6am in the morning, which doesn't happen normally in usual daily routine of the user. Figure 3 shows the detection of that unusual behaviour for day-10 in that specific time duration.

Figure 4 and Figure 5 shows the behaviour of the user for one month time duration. In week-1, the behaviour of the user remains average normal and this can be verified from Figure 2 that shows the regularity in the distributions of 'Home' and 'Work' patterns and shows that user did not make any unusual movements.

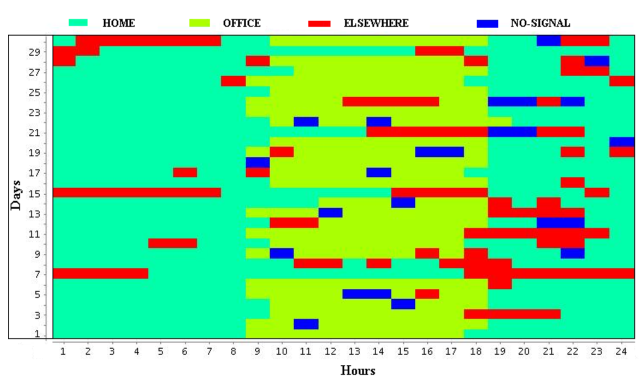

Figure 2. Distribution of (H, W, E and N) Transitions of Cell Tower ID Data

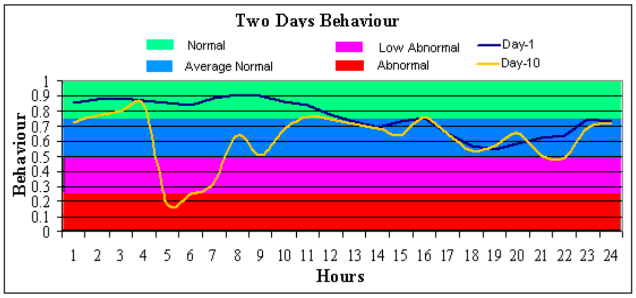

Figure 3. Comparison of Two Days of Behaviour Detected from Cell Tower ID Data

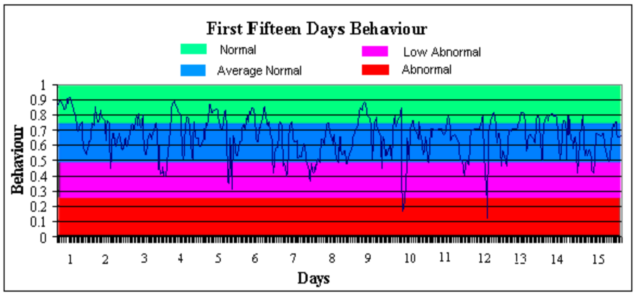

Figure 4. Week-1 and Week-2 Behaviour Detected from Cell Tower ID Data

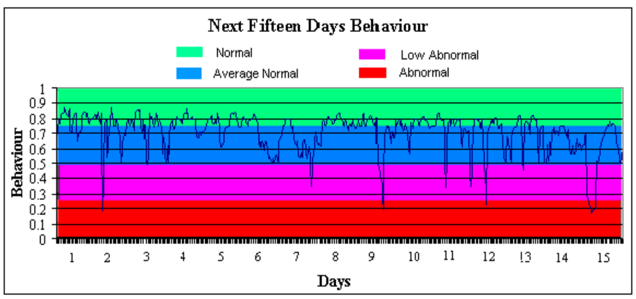

Figure 5. Week-3 and Week-4 Behaviour Detected from Cell Tower ID Data

In week-2, there is a change in behaviour on third and fifth day of the week when the user's (H, W, E and N) distributions show an irregular routine activity. In week-3 and week-4, some unusual routines are also detected in the behaviour of the user. The results presented above show only the cell tower ID data. We will now discuss the results obtained by using the same technique on only the Bluetooth proximity data of the same user. Each time slot for Bluetooth proximity data represents one hour as in the case of cell tower ID data.

Bluetooth proximity data is available in the form of detected devices as a result of a scanning performed by the user's cell phone after every five minutes. Each scanning results a list of devices present within the range of 5-10m. The first aim was to cluster the data in 'Home', 'Office' and 'Other Devices', so that the above technique can be used with Bluetooth data. The Bluetooth proximity data is clustered into only three categories so that the results obtained from cell tower ID data and the Bluetooth proximity data can be combine together to see if we could get some interesting anomalies in behaviour of the user. In future work, the aim is to cluster the Bluetooth proximity data into more finer grained time periods and try to detect the anomaly in user's behaviour in smaller time slots. The reason behind the clustering of Bluetooth data on finer scale is to classify the user behaviour in different activities that will provide one step further in the identification of unusual routines without using cell tower data.

After analysing the data, user's home computer device was given the name 'Home' (H). That means that all those time slots in which user detects his home computer device, considered as 'H' because it shows user's presence in the home. For office, there are many devices that user detect during office hours. To obtain a cluster of devices that belong to office, we remove the weekends from one month data and use Jaccard index [12], to detect how similar the detected devices are throughout the office hours for all remaining weekdays. Jaccard similarity equation is:

$$J(A,B)=\frac{|A\bigcap B|}{|A\bigcup B|}$$

where, 'A' and 'B' are sets of detected devices in two consecutive days. First, 'A' and 'B' represent day-1 and day-2, then day-2 and day-3 and so on up to the all remaining weekdays that left after removing the weekends from one month Bluetooth data. This gives us the similarity of detected devices during the office hours between the pairs of consecutive days, shown below in Figure 6. The average similarity between the detected devices is above 0.5. This means there are many devices that user detect repeatedly during his office hours. All those devices that user detect for at least 70% of the days during office hours goes in 'Office' cluster. All other devices go in 'Other Devices' cluster. After classifying the devices, a new data matrix is generated that contains twenty four time slots for each day as were in the case of cell tower ID data. Each time slot is assigned one of these clusters (i.e., Home, Office, Other Devices, No Devices Found) depending upon the number of detection of the devices belonging to a specific cluster. The same technique as used on cell tower IDs, described previously, is also used with Bluetooth proximity data. Figure 7 shows the Home/Work distribution of locations depending on the Bluetooth data clusters while Figure 8 shows the fifteen days behaviour of the user detected from both cell tower ID and Bluetooth proximity data. An interesting observation can be made by analysing the results of both cell tower ID and Bluetooth ID data. It is observed that sometimes when behaviour detected from cell tower ID data is not unusual, a change in behaviour is detected from Bluetooth proximity data. It can be said that it is more likely to be detecting unusual behaviour because during a regular routine of office hours of a weekday, user is supposed to detect 'Office Devices'. Cell tower ID data will show his normal behaviour as the user is in Office, but may be there is some gathering or meeting of students or staff that is not part of the regular routine. Behaviour detected from Bluetooth proximity data can be pointing towards that activity.

 ISBN: 978-1-61208-100-7

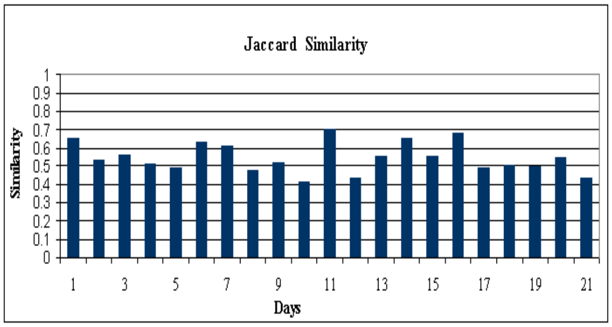

Figure 6. Jaccard Vertex Similarity

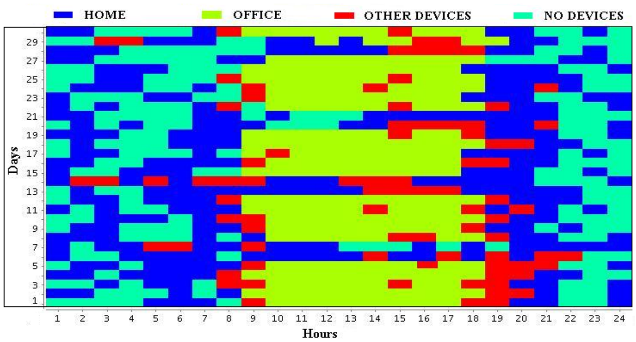

Figure 7. Distribution of Home/Work Transitions of Bluetooth Data

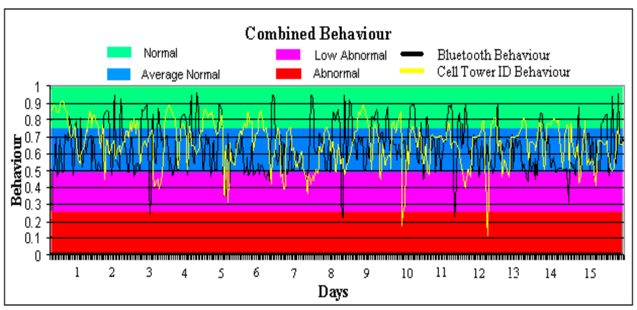

Figure 8. Fifteen Days Behaviour of Both Bluetooth and Cell Tower ID Data

These results show that for low entropy users, the detection of unusual routines and behaviours by using only Bluetooth data is possible. These low entropy users follow specific routines as compared to high entropy individuals, who live more diverse lives. This study aims to aid elderly people and patients to detect abnormal and unusual behaviours to avoid any accidents. Normally patients and elderly people have fixed and limited routines to follow that can likely be detected using Bluetooth devices by clustering the Bluetooth device detections into different activities or communities.

V. SUMMARY AND FUTURE WORK

In this paper, real time Bluetooth proximity and cell tower ID data is used to detect abnormal and unusual activities and routines of an individual by using neural networks. A low entropy user was selected for experiments due to the regularity and constancy in his routines. A successful detection of abnormal behaviour in this user's routines is done by using cell tower ID's and Bluetooth proximity data. Bluetooth proximity data is only clustered into three different categories. The idea was to combine the results of behaviour detection from Bluetooth proximity data with the results of cell tower ID data. To detect anomalies in more specific and lower level activities and routines, we need to cluster the Bluetooth proximity data into temporal clusters. In future work, we will try to cluster the Bluetooth proximity data on temporal scale to cover the minute details of the user's behaviour and will also try to predict the behaviour based on these clusters and communities detection. This will help us to facilitate elderly people and patients who need more care and concern about their behaviour and unusual routines that can cause serious accidents.

REFERENCES

[1] K. Farrahi and D. Gatica-Perez, "Daily Routine Classification from Mobile Phone Data". In: Popescu-Belis, A., Stiefelhagen, R. (eds.) MLMI 2008. LNCS, vol. 5237, pp. 173–184. Springer, Heidelberg (2008)

[2] K. Farrahi and D. Gatica-Perez, "What did you do today? Discovering daily routines from Large-Scale Mobile Data". In: MM 2008: Proceeding of the 16th ACM international conference on Multimedia, pp. 849–852. ACM, New York (2008)

[3] M. Hermersdorf, H Nyholm, J Perkiö, and V Tuulos. "Sensing in Rich Bluetooth Environments"- Workshop on WorldSensorWeb, in Proc. SenSys, 2006 - sensorplanet.org

[4] N. Eagle and A. Pentland, "Reality mining: sensing complex social systems". Personal and Ubiquitous Computing 2006 – Springer, Vol. 10, # 4, 255-268

[5] GPS enabled tracking devices, retrieved on 14-07-2010, http://www.elderoptionsoftexas.com/article_alzheimers_gps_tracking_devices.htm

[6] M. Vukovic, G. Vujnovic, and D. Grubisic, "Adaptive User Movement Prediction for Advanced Location-aware Services", Proceedings of the 17th international conference on Software, Telecommunications and Computer Networks, pp. 343-347 (2009)

[7] M. Vukovic, I. Lovrek, and D. Jevtic. "Predicting user movement for advanced location-aware services". In 15th International Conference on Software, Telecommunications and Computer Networks, pages 1–5. SoftCOM 2007., 2007.

[8] N. Eagle, "Machine Perception and Learning of Complex Social Systems", Ph.D. Thesis, Program in Media Arts and Sciences, MIT, June 2005

[9] M. A. Azam and T. Laurissa (2009). Behaviour Detection Using Bluetooth Proximity Data. Proceedings of Networking & Electronic Commerce Research Conference pp. 46-52 (NAEC 2009).

[10] K. Hara, T. Omori, and R. Ueno, "Detection of unusual human behaviour in intelligent house"; Proceedings of the 2002 12th IEEE workshop on Neural Networks for Signal Processing, pp. 697-706, 2002.

[11] Linear Neural Networks, retrieved on 14-07-2010, retrieved from http://www.idsia.ch/NNcourse/linear2.html

[12] Jaccard Index, retrieved on 14-07-2010, retrieved from http://www.statemaster.com/encyclopedia/Jaccard-index

[13] T. Yiping, Z. Zhiying, G. Hui, L. Huiqiang, W. Wei, and X. Gang, "Elder Abnormal Activity Detection by Data Mining", SICE Annual Conference in Sapporo, August 4-6, 2004, vol. 1, pp. 837–840 (2004) Japan

[14] N. Eagle, A. Clauset, and J. A. Quinn, "Location Segmentaion, Inference and Prediction for Anticipatory Computing", AAAI Spring Symposium, 2009

[15] D. H. Hu, X. Zhang, J. Yin, V.W. Zheng, and Q. Yang, "Abnormal Activity Recognition Based on HDP-HMM Models", AAAI Publications, 21st International Conference on Artificial Intelligence, pp. 1715–1720, 2009

[16] C. Wren, Y. Ivanov, I. Kaur, D. Leigh, and J. Westhues, "SocialMotion: Measuring the Hidden Social Life of a Building". In: J. Hightower, B. Schiele, and T. Strang, (eds.) LoCA 2007. LNCS, vol. 4718, pp. 85–102. Springer, Heidelberg (2007)

[17] I. McCowan,, D. Gatica-Perez, S. Bengio, and G. Lathoud,.: Automatic Analysis of Multimodal Group Actions in Meetings. IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI) 27(3), 305–317 (2005)

[18] R. Stiefelhagen, K. Bernardin, H.K Ekenel, J. McDonough,, K. Nickel, M. Voit, and M. Woelfel: Audio-Visual Perception of a Lecturer in a Smart Seminar Room. In: Signal Processing - Special Issue on Multimodal Interfaces, vol. 86 (12). Elsevier, Amsterdam (2006)

[19] T. Nicolai, N. Behrens, and E. Yoneki, "Wireless Rope: Experiment in social proximity sensing with Bluetooth". In *fourth annual IEEE International Conference on Pervasive Computing, LNCS 4277, 2006*

[20] T. M. Gill, M.M. Desal, A. Evelyne, T. R. Holford, and C. S. Williams, "Restricted-Activity among Community-Living Older Persons: Incidence, Participants, and Health Care Utilization", Annals of Internal Medicine. http://www.annals.org/content/135/5/313.full.pdf+html

 ISBN: 978-1-61208-100-7