Sensing Room Occupancy Levels with Signal-to-Noise Ratio and Signal Phase and Multiple Antenna Configurations

Jens Weppner
and Benjamin Bischke
DFKI
Kaiserslautern, Germany, 67663

Attila Reiss
and Robert Duerichen
Robert Bosch GmbH, Corporate Research
Renningen, Germany, 71272

Paul Lukowicz
DFKI and
Technische Universität Kaiserslautern
Kaiserslautern, Germany, 67663

***Abstract*—We propose a novel machine learning based method for estimating the number of people present in a room (e.g. in a shared office space) based on WiFi signal-to-noise ratio and signal phase data provided by WiFi Channel State Information compatible hardware. We apply random decision forests machine learning and show that the precise number of people can be estimated with a score of 0.66 and the occupancy levels (empty, low, high) with a score of 0.87 at an affordable cost. We evaluate our approach in two settings: one small room with 0–2 and in a medium-sized office space with 0-8 people performing their usual office desk work. Beyond determining maximum recognition rates we systematically investigate the impact of different design choices (antennas, training data) on system performance. The proposed method outperforms a statistical baseline method significantly.**

I. Introduction

Knowing room occupancy levels (how many people are in a given room) is important for a number of applications. Examples range from evacuation coordination in emergency situations through the optimization of office space usage to energy management. Given the fact that today WiFi can be found in nearly all buildings and public spaces, using the disturbance that people cause, WiFi signals as a virtual sensor is an attractive approach for occupancy level estimation. However, to date, WiFi signal analysis has mostly been used for estimation of qualitative crowd density (which is associated with strong variation in the signal strength, see related work section) and for the detection of movement (which cause characteristic temporal signal fluctuations, see related work section). The estimation of the number of people in a small group (≤ 10) who is largely static (e.g. working at their desks) in a shared space such as an office is a more difficult problem as it is related to subtle, mostly static signal changes. To address this problem we present a new sensing concept that applies machine-learning techniques to appropriate features extracted from the Channel State Information (CSI, [1], [2]) data.

II. Related Work and Paper Contributions

Different sensors have been proposed to estimate the occupancy level of spaces. These include surveillance cameras [3], thermal sensors [4], pressure sensors [5], acoustic sensors [6] and floor pressure sensors [7]. Compared with most of the above sensing modalities WiFi signals ubiquitous in most environments. Work so far includes strategies evaluating subjects beacons (i.e. smartphones) for crowd density estimation by collaborative mobile sensors [8] by citywide mobile participatory sensors [9] or by stationary sensors [10] based on device identifier and received signal strength indicator (RSSI) features. RSSI methods are suitable for qualitative crowd density estimation, indoor localization and many other interesting applications, but its limits arise when subtle signal variations due to few mostly static people being present need to be measured. The work most similar to ours was done by Depatla et al. [11] and Xi et.al [12]. However, having in common to build on top of CSI towards crowd counting, their objective and method is different. In [11] they focus on a continuously walking crowd and build a mathematical motion model based on RSSI, compared to the CSI approach presented in this paper. In [12] they build on continuously moving people and apply a logistic function and deploy a dense grid of sensors in the environment. Our objective is going one step further by defining a method for occupancy level estimation in a challenging environment with mostly static/sitting people.

In this paper we present a WiFi-CSI scan-based method to estimate the fill level within a shared office space. The main contributions are as follows:

1 We propose a method to estimate the exact number of people and the occupancy level (empty, low, high) in office rooms while measuring WiFi signal-to-noise ratio and signal phase with different kinds of antenna setups.
2 We present two experimental evaluations on a small office room and a medium-sized shared office space.

III. Methods

A. Background

The interaction of WiFi signals with the environment is a complex process that involves absorption, reflections (multi-path) and a variety of wave specific effects (refraction, interference etc.). When looking at phenomena that are determined by

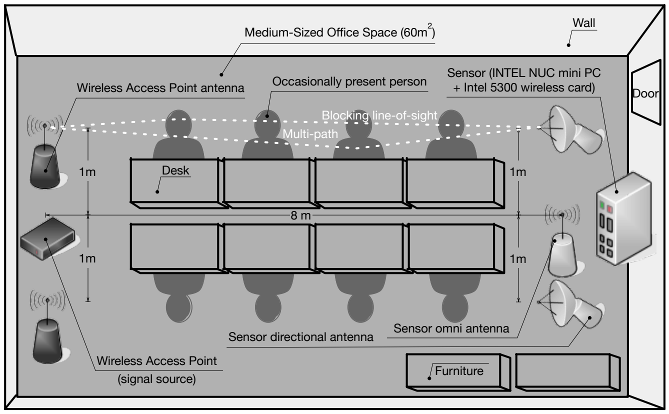

Fig. 1: Experimental environment. Wireless signals are influenced due to multiple effects, i.e. *blocking in the line-of-sight* and *multi-paths* due to objects as well as human bodies. The challenge is to measure and learn from such transient signal variations.

high degree of signal blockage (e.g. detecting a dense crowd) much of the complexity can be ignored as the received signal strength (given by RSSI) can be used for analysis. However, when considering subtle influences caused by a small number of largely static people a more complex metric is needed. In the IEEE 802.11n standard such a metric is provided by Channel State Information that captures signal strength and phase information for Orthogonal Frequency Division Multiplexing (OFDM) subcarriers and between each pair of transmit-receive antennas. It has originally been defined to allow the sender to improve the link via transmit beam-forming (see [1]).

B. Experimental Environment

One small-sized ($20m^2$) and one medium-sized ($60m^2$) sized shared office space were selected. In the small office 0-2 people were present, and in the medium office space 0-8 people were present doing their work at desks as usual. The omni-directional antenna of the sensor was placed at one side of the room and the directional antennas of the sensor were positioned with a 1m sideways displacement to the omni-directional antenna at line of sight between the desk locations and each sender antenna (see Figure 1).

C. Ground Truth

Collecting ground truth was a multi-step procedure. Firstly, we installed a video camera in accordance to the people during the experiment. Secondly, after the experiment the camera images were manually annotated each 5 seconds with the exact seat position of each person. Thirdly, the total number of people was derived from the individual person annotations.

D. Hardware

We configured a prototype sensing unit based on an INTEL NUC mini PC running Ubuntu with a MIMO Intel 5300 card with customized external antenna connections. We connected one omni-directional antenna and two displaced (1m) directional antennas. As a signal source we used a conventional MIMO WiFi access point with 1m displaced antennas.

E. Measurement and Features Computation

To precisely measure the phase information the sender and receiver must be perfectly synchronized. Unfortunately, commercial WiFi devices have non-negligible carrier frequency offsets. Nevertheless we can identify and calculate the signal phase difference by the two signal streams coming from both sender antennas synchronously. The signal phase difference is calculated for each of the 30 subcarriers (based on OFDM) and for each sensor-antenna plus dual-sender antenna group. Based on the phase difference we define the mean, variance, minimum and maximum antenna group signal phase difference within a 1 second time windows. The signal-to-noise ratio (SNR) is extracted for each pair of sender and sensor antenna (in total six SNR measurements per carrier). Based on the SNR we defined the features mean, variance, minimum and maximum SNR within 1 second time windows. For each sensor antenna the total number of features is 360 (2*40*30 (SNR) + 4*30 (phase)). We observed that changes in the window parameter (from 1 to 10 seconds in 1 second steps) have insignificant influence to the classification result ($\pm 2\%$ F1-Score). However smaller window sizes require significantly more classifier training time.

F. Machine Learning Based Fingerprinting

Each people arrangement, according to ground truth, is used as a fingerprint of the people count. 70% of the fingerprint were either used in the training step and 30% was used as the test set. Together with the ground truth targets we trained random decision forests (10 decision trees) machine learning classifier. We also compared the results to other classifiers such as SVM and ensemble RandomForest classifier resulting in similar results but are up to 14 times slower during training time.

IV. Evaluation

V. Results

A. Small-Sized Office Space ($20m^2$)

We evaluated the 3-class classifier with the CSI approach with signal-to-noise ratio and the signal phase difference features with a f1-score of 0.73 while classifying between 0, 1 or 2 persons in the room (see Table I). For comparison, we evaluated the classifier with the traditional signal strength approach with a score of 0.50.

Antennas	**Features**	**Classification results**		
		F1-Score	Precision	Recall
ALL	RSSI	0.50	0.50	0.51
ALL	SNR&phase	0.73	0.73	0.73

TABLE I: Small office (20 m^2) experiment results. SNR and phase method in comparison with traditional RSSI method. Exact classification between 0, 1 or 2 persons in room.

B. Medium-Sized Office Space ($60m^2$)

We compared a baseline pure statistical hourly knowledge, derived from the ground truth, to the the 9-class classifier, while estimating 0,1,2,3,4,5,6,7 or 8 persons in the shared office space. Our reference method has a score of 0.27 and our new method has a score of 0.67 (see Table II). The results are inferior compared to a simple hourly statistical evaluation which is due to fluctuations of office space fill level over days.

The main objective of this research is considering different levels of complexity of the sensor setup. Our method based on the sensor information of just a single omni-directional antenna results in a score of 0.61, with two carefully placed directional antennas a score of 0.62, and with the combination of omni-directional directed antennas in a score of 0.67 (see Table III for detailed scores).

Due to background knowledge of the actual need for occupancy level granularity (empty, low, high occupancy) we derived a 3-class classification problem from the ground truth information (0, 1-3 or 4-8 persons). This results in score of 0.87 (see Table IV).

Another important objective is the amount of data used in machine learning for training the classifier as labeled training data is costly to acquire. In Figure 2 we visualize the amount of training data from 10% to 100%, which is always a subset of the whole data set, together with the resulting score. With the 9-class problem (exact people count) we see a nearly continuous increase in the score with increasing training data. With the simplified 3-class problem (0, 1-3, 4-8 persons) we see a plateau once we reach 50% of the training data available.

Approach	**Classification Results**		
	F1-Score	Precision	Recall
Baseline statistics	0.26	0.27	0.26
SNR&PHASE, all antennas	0.67	0.66	0.66

TABLE II: Medium office ($60m^2$) experiment results on baseline statistical classifier sampling the number of persons per class from a distribution created for every hour (distribution was estimated from the Ground Truth).

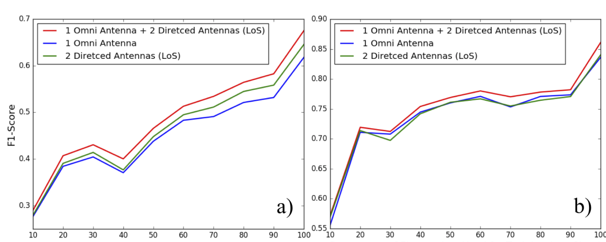

Fig. 2: Relation between the amount of labeled training data (costly to acquire) and the classification score. a) 9-class classification (exact people count). b) 3-class classification (empty, low, high occupancy level).

Antennas	**Feature Sets**	**Classification results**		
		F1-Score	Precision	Recall
OMNI	SNR	0.59	0.59	0.59
	PHASE	0.51	0.52	0.52
	SNR&PHASE	0.61	0.61	0.60
DIRECTIONAL	SNR	0.61	0.61	0.61
	PHASE	0.52	0.53	0.52
	SNR&PHASE	0.62	0.62	0.62
ALL	SNR	0.65	0.65	0.65
	PHASE	0.58	0.58	0.57
	SNR&PHASE	0.67	0.66	0.66

TABLE III: Medium office ($60m2$) experiment results. Comparison between different antenna setups and signal characteristics: 1 omni-directional antenna, 2 directional antennas, or 1 omni-directional + 2 directional antennas. Using multiple features on signal characteristics of signal-to-noise (SNR) and signal phase (PHASE) measurements. Classification between 0,1,2,3,4,5,6,7,8 persons in shared office space.

Class Range	**Classification results**		
	F1-Score	Precision	Recall
0,1,2,3,4,5,6,7,8	0.67	0.66	0.66
0,1–3,4–8	0.87	0.87	0.87

TABLE IV: Medium office ($60m^2$) experiment results. Same setup as in Table III but in comparison to classification on 0,1–3 and 4–8 persons in shared office space.

VI. Conclusions and Future Work

We presented a challenging problem in occupancy level estimation, a method and the evaluation. We outperformed both the statistical baseline method and the traditional signal strength approach. We presented results on the exact people count and on the occupancy level (empty, low, high). While additional antennas results in an increased estimation score, the amount of training data is the most significant factor. Future work will focus on deep learning and time series based estimation approaches to enhance the estimation. The data set and ground truth annotation will be made available to the public.

References

[1] "Ieee std. 802.11n-2009: Enhancements for higher throughput," http://www.ieee802.org, last Checked: April 2017.

[2] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: gathering 802.11 n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1. IEEE, 2005, pp. 886–893.

[4] F. Wahl, M. Milenkovic, and O. Amft, "A distributed pir-based approach for estimating people count in office environments," in *Computational Science and Engineering (CSE), 2012 IEEE 15th International Conference on.* IEEE, 2012, pp. 640–647.

[5] J. Cheng, B. Zhou, M. Sundholm, and P. Lukowicz, "Smart chair: What can simple pressure sensors under the chairs' legs tell us about user activity," in *UBICOMM13: The Seventh International Conference*

on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2013.

[6] P. G. Kannan, S. P. Venkatagiri, M. C. Chan, A. L. Ananda, and L.-S. Peh, "Low cost crowd counting using audio tones," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '12. New York, NY, USA: ACM, 2012, pp. 155–168.

[7] T. Mori, Y. Suemasu, H. Noguchi, and T. Sato, "Multiple people tracking by integrating distributed floor pressure sensors and rfid system," in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, vol. 6. IEEE, 2004, pp. 5271–5278.

[8] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *Proceedings of the Eleventh Annual IEEE International Conference on Pervasive Computing and Communications (Percom 2013)*. IEEE, 3 2013, pp. 193–200.

[9] J. Weppner, P. Lukowicz, U. Blanke, and G. Troster, "Participatory bluetooth scans serving as urban crowd probes," *Sensors Journal, IEEE*, vol. 14, no. 12, pp. 4196–4206, Dec 2014.

[10] J. Weppner, B. Bischke, and P. Lukowicz, "Monitoring crowd condition in public spaces by tracking mobile consumer devices with wifi interface," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp-16), September 12-16, Heidelberg, Germany*, ser. UbiComp '16. ACM, 2016, pp. 1363–1371.

[11] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only wifi power measurements," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1381–1393, 2015.

[12] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using wifi," in *INFOCOM, 2014 Proceedings IEEE*, April 2014, pp. 361–369.