A Trained-once Crowd Counting Method Using Differential WiFi Channel State Information

Conference Paper · June 2016

DOI: 10.1145/2935651.2935657

4 authors, including:

Simone Di Domenico
University of Rome Tor Vergata

Mauro De Sanctis
University of Rome Tor Vergata

Ernestina Cianca
University of Rome Tor Vergata

A Trained-once Crowd Counting Method Using Differential WiFi Channel State Information

Simone Di Domenico Mauro De Sanctis Ernestina Cianca

Giuseppe Bianchi
University of Roma Tor Vergata, via del politecnico 1, 00133, Rome, Italy
{simone.di.domenico; mauro.de.sanctis; ernestina.cianca; giuseppe.bianchi}@uniroma2.it

ABSTRACT

This paper focuses on the problem of providing a rough count of the number of people in a room using passive WiFi Channel State Information (CSI) measurements taken by a *single* commodity receiver. The feature which mainly distinguishes our work from others is the attempt to emerge with an approach which *does not require* any dedicated training inside the specific environment where the system is deployed. Our proposal stems from the intuitive observation that features which account for *variations* of CSI are expected to be less sensitive to the surrounding environment as opposed to features which account for *absolute* CSI measurements. We turn such intuition into a concrete proposal, by suitably identifying a set of *differential* CSI feature candidates, and by selecting the (two) most effective ones via minimization of the summation of the Davies-Bouldin indexes. We preliminary assess the effectiveness of the proposed approach by training *once for all* the system in a room, and testing the system in two *different* rooms having different size and furniture, and involving people freely moving in the rooms with no a-priori movement constraints.

Keywords

Device-free, People Counting, WiFi, RF sensing, Crowd Counting, CSI, RSSI.

1. INTRODUCTION

Crowd counting is the process of estimating the number of people in a given area, in either an open or closed environment. This process may be required in both safety critical and uncritical situations. Tracking human queues or counting people sitting in a waiting room in areas like retail stores or airports, hospitals and theme parks [1] could be used for improving the offered services. Such improved services result in a better user satisfaction but also reduced costs and improved efficiency. In the transportation, bus and train schedules or boarding and payment processes could be adjusted according to the number of people waiting for the service.

In most of the applications, the required accuracy of this estimation decreases as the number of people increases. On one hand, it is important to distinguish between empty room, one person in a room, two persons in a room. On the other hand, for a higher number of people it is sufficient to have a density estimation like stating what is the probability that there are, say, 3 to 5 people, or 6-10, or 11-20, and so on.

In this paper we focus on RF-based device-free sensing which does not require any specific user cooperation [2]. RF-based device-free crowd counting techniques are in general based on some kind of measurement of the impact of the presence of different numbers of people on the RF propagation channel. So far, most of the approaches are based on measuring and properly processing the RSSI of the RF signal [3], [4], [5], [6], [7]. A more recent approach uses the Channel State Information (CSI) measurements available with a specific tool for the Intel 5300 WiFi card to describe the propagation channel variations [8], [9].

It is worth noting that all these works requires a training phase, which *must be performed in the same environment* as the one in which the crowd counting is needed, with the same setting conditions (e.g., same position of transmitters and receivers). So, these techniques cannot be used in environments different from the ones in which the training phase has been performed. This is the biggest limit of RF-based device-free crowd counting and more generally of activity recognition technique. While it is recognized that overcoming this limit is fundamental, to the best of our knowledge very little work has been spent so far on such more challenging problem. This paper documents our first step in this important direction.

Specifically, in this paper, we propose and preliminarily assess a crowd counting method that is based on differential measurements of the CSI gathered by a single WiFi receiver and a normalization process. The usage of *differential* CSI features (opposed to *absolute* CSI features) stems from the intuitive fact that the multipath experienced in different environments of course significantly varies from room to room, depending on size, topology, furnishing, materials, and so on. But the *difference* in the multipath experienced by consecutive frames is arguably less sensitive to the (static) environment characteristics and instead is expected to depend mainly from the number of people moving in the environment.

WPA '16, June 26, 2016, Singapore.

DOI: http://dx.doi.org/10.1145/2935651.2935657

In the remainder of this paper we identify specific features related to CSI variations, we propose a relevant normalization, and we apply machine learning techniques (specifically, minimization of the sum of the Davies-Bouldin indexes) to select the most effective pair of features. Furthermore, we show, through experimental results, that this method can be used in environments different than the ones over which the training phase is performed. The achieved performance in terms of accuracy/error are promising, also in comparison with the performance of other techniques proposed in literature, or the same technique but used after a training phase performed exactly in the same environment where the crowd counting is performed. Moreover, these promising performance have been achieved even in tests conditions more difficult than the ones considered in other previous works. More specifically:

- a single WiFi Access Point (AP) and a single WiFi receiver are used whilst most of the related works use multiple APs/trasnmitters and receivers;
- volunteers move as they wanted (randomly, repeating a path, etc.) or they simply stand;
- we have set a time duration of the rolling window $W = 10$ s that is an acceptable value for the estimation delay. A larger window size allows to increase the accuracy thanks to a better statistical analysis, however, a large window size is helpless when the crowd size changes more frequently than the estimation delay.

2. ARCHITECTURE AND EXPERIMENTAL SETUP

In this Section, we present the architecture and the experimental setup of our crowd counting system, which is based on CSI data collected from WiFi AP. The system includes the following modules: a commercial standard 2.4 GHz WiFi b/g/n AP which acts as a double antenna transmitter; a laptop with Ubuntu 10.04 LTS with the Intel WiFi Link 5300 wireless NIC with 3 antennas which acts as receiver. A Matlab script running on the same laptop processes offline the collected data and provides the information about the estimated number of people. The laptop sends ICMP echo request packets every 20 ms to the WiFi AP and receives the ICMP echo reply packets.

CSI data are extracted from the received packets by using a modified version of a open-source wireless driver [9]. We carried out WiFi CSI collection on three different rooms, see Figure 1:

- Room A: small size office room (5m × 6m).
- Room B: medium size meeting room (5m × 9m).
- Room C: large size meeting room (6m × 12.5m).

The furniture and the location of the AP and the WiFi receiver is shown in Figure 1. The distance between the AP and the WiFi receiver was larger than 5 m in any room, while the distance between any volunteer and the AP (or the receiver) was larger than 1.5 m. The size of the crowd was limited to $N_c = 7$ people. No guidelines or limitations to movements have been given to the volunteers which could walk or stand firm following a random sequence. A number of $N_{CSI} = 5000$ CSI measurements has been collected for

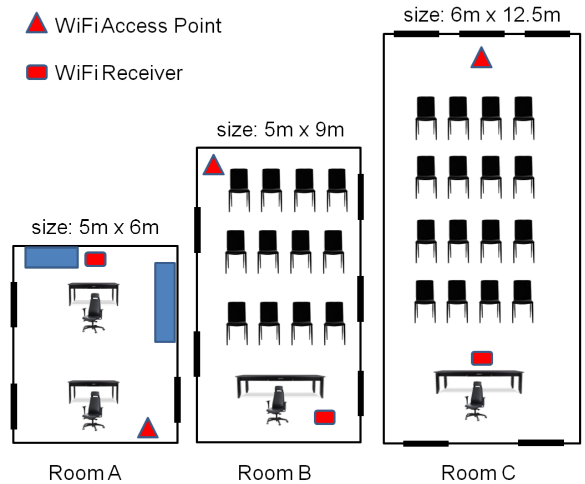

Figure 1: Experimental environments.

each class of people (from 0 to $N_c = 7$) and for each room. The collected CSI data are processed offline to build the dataset performing data fusion, feature extraction, feature selection and classification tasks.

3. DIFFERENTIAL APPROACH

In the following Section, we describe the methodology of our CSI-based crowd counting system. First, we present the source data, i.e. the CSI, exploited by the counting system. After that, we show which features have been considered and then selected to maximize the performance. Lastly, we briefly describe the classification algorithm.

3.1 CSI data

In radio communications, the transmitted electromagnetic waves reach the receiving antenna by more paths, this phenomenon is also known as multipath propagation. The presence and the movements of a human body imprint an information on the intercepted multipath components and hence on the Channel Frequency Response (CFR), making it possible the recognition of its presence, activity and gesture without wearing any device or sensor.

CFR variations are exploited by our system, in particular, in place of CFR we use the CSI estimated by the WiFi OFDM receiver. An OFDM-based transmission encodes the digital data stream on a set of orthogonal sub-carriers. Let us denote with $\mathbf{y}$ the received complex vector after the FFT at the receiver, $\mathbf{y}$ can be expressed as follows:

$$\mathbf{y} = \mathbf{h_c} \cdot \mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ is the transmitted complex vector, and $\mathbf{h_c}$ is the complex channel response vector in the frequency domain, and $\mathbf{n}$ is the additive gaussian noise vector. The *i-th* component of $\mathbf{h_c}$ is the CSI value of the i-th OFDM subcarrier, which represents the complex channel gain for the *i-th* subcarrier. To estimate $\mathbf{x}$, the receiver needs to estimate the channel vector $\mathbf{h_c}$, so this estimation is available at the receiver. Therefore, we used the modified firmware of the IWL 5300 NIC to gather the raw CSI values available at the receiver. The used firmware provides the CSI value for $N_{sub} = 30$ OFDM subcarriers. The complex vector containing the 30 CSI values (hereinafter CSI vector) can be expressed as

$$\mathbf{h} = [h_1, h_2, ..., h_{30}], \tag{2}$$

where the CSI of the i-th subcarrier is defined as:

$$h_i = |h_i| e^{j\angle h_i}, \tag{3}$$

where $|h_i|$ is the amplitude and $\angle h_i$ is the phase of the channel of the *i-th* subcarrier. In our experiments we have used 2 transmit antennas and 3 receive antennas, therefore $N_{ch} = 6$ RF channels are available for processing. This means that for each received packet N_{ch} CSI vectors can be used to extract the desired features.

3.2 CSI feature extraction

The key idea behind the choice of the features that can make the system performance as independent as possible from the need of a training phase is the following: when there are no people in the room, the multipath environment is nearly constant and the observed variations are only due to the thermal noise and interference. On the other hand, if one person is moving inside the room, one or more multipath components change over the time, consequently, the CSI across all the subcarriers varies. Moreover, it is intuitive that two people walking into the room introduce higher multipath variation compared to one person walking. In general, an increase of the number of people also increase the multipath variations as a consequence of the higher number of moving scatterers. Therefore, our crowd counting system is based on a measurement of how the CSI varies over the time, which should depend mainly from the number of people moving in the environment rather than the characteristic of the environment itself (like the size, the furniture, the material of walls, etc.).

Figure 2 shows a set of collected CSI curves for the empty, one person, two persons and three persons scenarios. This Figure supports our intuition, in fact, it shows that the more are the persons in the environment, the more is the CSI variation across all the subcarriers.

In particular, we have chosen to use as a basic descriptor the Euclidean distance between two CSI vectors, i.e between two CSI. Two important steps of the feature extraction are the proper choice of the normalization to be performed, which can further help in reducing the dependence from the background environment, and also the selection of the scheme to combine the information collected from the N_{ch} RF channels. The following normalizations have been considered:

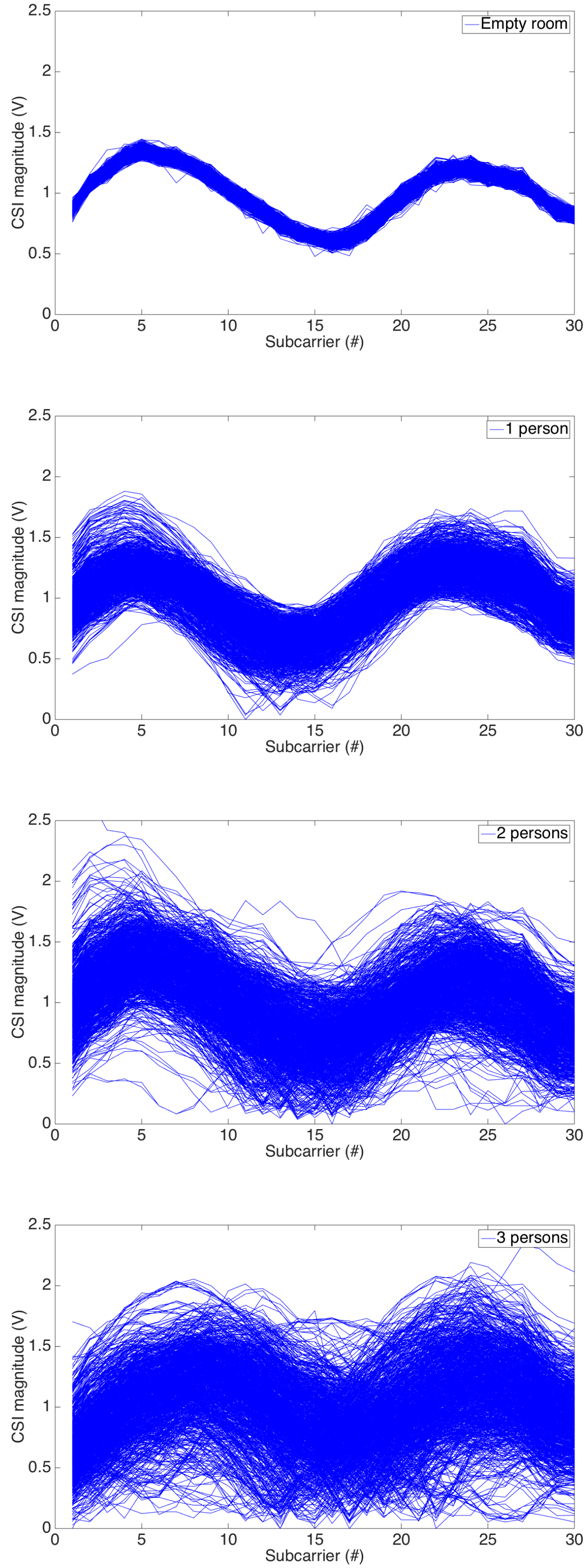

Figure 2: CSI curves for different scenarios

$$\hat{\mathbf{h}}_j^k = \begin{cases} \mathbf{h}_j^k & \text{(4a)} \\ \dfrac{\mathbf{h}_j^k}{m(\mathbf{h}_j^k)} & \text{(4b)} \\ \dfrac{\mathbf{h}_j^k}{\sqrt{e(\mathbf{h}_j^k)}} & \text{(4c)} \end{cases}$$

where $\mathbf{h}_j^k$ is the CSI vector of the received packet at the *k-th* time sample on the *j-th* channel, and $m(\mathbf{h}_j^k)$ and $e(\mathbf{h}_j^k)$ are the sample mean and the sample energy of the $\mathbf{h}_j^k$ respectively.

The second level of normalization is related to the Euclidean distance computed between two CSIs. In this case, the following normalizations have been considered:

$$d_j^{k,l} = \begin{cases} d_{euc}(\mathbf{h}_j^k, \mathbf{h}_j^l) & \text{(5a)} \\ \dfrac{d_{euc}(\mathbf{h}_j^k, \mathbf{h}_j^l)}{m(\mathbf{h}_j^k) \cdot m(\mathbf{h}_j^l)} & \text{(5b)} \\ \dfrac{d_{euc}(\mathbf{h}_j^k, \mathbf{h}_j^l)}{\sqrt{m(\mathbf{h}_j^k) \cdot m(\mathbf{h}_j^l)}} & \text{(5c)} \\ \dfrac{d_{euc}(\mathbf{h}_j^k, \mathbf{h}_j^l)}{\sqrt[4]{\Gamma^k \cdot \Gamma^l}} & \text{(5d)} \end{cases}$$

where the function $d_{euc}(\mathbf{h}_j^k, \mathbf{h}_j^l)$ is the Euclidean distance between the *k-th* and *l-th* CSI vectors received on the *j-th* channel, and Γ^k and Γ^l are the total received signal strength for the *k-th* and *l-th* received packet respectively. The distance between CSIs has been computed for each channel, therefore N_{ch} sets of distances were available for the feature extraction.

The distances computed over the N_{ch} channels have to be fused together and to this end we have considered the following combining schemes:

$$D^{k,l} = \begin{cases} \displaystyle\sum_{j=1}^{N_{ch}} d_j^{k,l} & \text{(6a)} \\ \displaystyle\sum_{j=1}^{N_{ch}} \frac{d_j^{k,l}}{\sqrt[4]{\Gamma_j^k \cdot \Gamma_j^l}} & \text{(6b)} \\ \displaystyle\sum_{j=1}^{N_{ch}} \frac{\Gamma_j^k \cdot \Gamma_j^l}{\Gamma^k \cdot \Gamma^l} \cdot d_j^{k,l} & \text{(6c)} \end{cases}$$

where Γ_j^k and Γ_j^l are the received signal strength for the k-th and l-th received packet on the j-th channel respectively. Notice that $\Gamma_1^k = \Gamma_4^k$, $\Gamma_2^k = \Gamma_5^k$, and $\Gamma_3^k = \Gamma_6^k$. This is because these couples of channels share the same receive antenna, thus the received signal strength indicator is the same.

The last step to compute the features is applying a statistical analysis of the combined distance over a temporal sliding window of size W packets. The width W of the rolling window has an important role on the classification accuracy of the crowd-counting system. In fact, increasing the window size allows to collect and process more data, improving the performance, but at the same time also increasing the estimation delay. The statistical moments have been evaluated using both the distance between consecutive CSIs and between all pairs of CSIs within the sliding window. Considering all the distances within a temporal window can lead to extract features that are more informative for the crowd density recognition. The statistical analysis was performed through the first and second order statistical moments, so the extracted features are the following:

$$F \begin{cases} m(D) & \text{(7a)} \\ \sigma(D) & \text{(7b)} \\ \dfrac{\sigma(D)}{m(D)} & \text{(7c)} \\ \dfrac{\sigma^2(D)}{m(D)} & \text{(7d)} \end{cases}$$

where $m(\cdot)$ is the sample mean, $\sigma(\cdot)$ is the sample standard deviation, $\dfrac{\sigma(\cdot)}{m(\cdot)}$ is the sample coefficient of variation, and $\dfrac{\sigma^2(\cdot)}{m(\cdot)}$ is the sample index of dispersion.

3.3 Feature Selection and Classification

In our approach we used a pair of features selected from the overall set of features previously defined. In order to select the best combination of feature which maximizes the classification accuracy we computed the Davies-Bouldin index $DB_{p,q}$ for a pair of classes (p, q) which is given by:

$$DB_{p,q} = \frac{||\boldsymbol{\sigma_p}|| + ||\boldsymbol{\sigma_q}||}{||\boldsymbol{\mu_p} - \boldsymbol{\mu_q}||} \tag{8}$$

where $\boldsymbol{\mu_p}$ and $\boldsymbol{\sigma_p}$ are the sample mean and the sample variance of a pair of features (F_a, F_b) for the p-th class. In addition, $p, q = 0, 1, 2, ..., N_c$, $DB_{p,q} = DB_{q,p}$ and $DB_{p,p} = 0$.

Then, we applied an exhaustive search optimization method over all pairs of features where the objective function $f(F_a, F_b)$ to minimize is the summation of Davies-Bouldin indexes:

$$f(F_a, F_b) = \sum_{p=0}^{N_c - 1} \sum_{q=p+1}^{N_c} DB_{p,q} \tag{9}$$

In this optimization step, the three datasets of CSI measurements collected in the three rooms have been merged. Therefore, the classes regarding the same number of people are merged for the three rooms, and we aim to minimize the intra-class distances and to maximize the inter-class distances building well separated and compact clusters based on the selected pair of features.

Finally, we found that the best pair of features is represented by:

1. F_a: F (7a), $D^{k,l}$ (6a), $d_j^{k,l}$ (5d), $\hat{\mathbf{h}}_j^k$ (4b).
2. F_b: F (7d), $D^{k,l}$ (6a), $d_j^{k,l}$ (5d), $\hat{\mathbf{h}}_j^k$ (4b).

The estimation of density or the estimation of the number of people was achieved using the linear discriminant classifier.

4. EXPERIMENTAL RESULTS

The classifier was trained using 50% of the dataset collected in Room B; Room B was chosen to train the classifier since it represents the middle size room with respect to Room A and Room C. Classification results are presented in the following subsection. Classification tests in Room A and Room C are performed using the overall datasets of the corresponding room, while classification tests in Room B are performed using 50% of the dataset of Room B.

Results are presented in Table 1 and 2 in terms of confusion matrices for the crowd counting case (8 classes) and for the occupancy estimation case (5 classes).

For the crowd counting case, the results show that more than 91% and 81% estimation errors are less or equal than 2 persons in Room A and Room C respectively. Good performance in terms of average accuracy is achieved in Room A, while acceptable average accuracy is achieved in Room C.

It is worth noting that in the smallest room, i.e Room A, the classification accuracy for a small number of persons is higher than that obtained in the biggest room, i.e Room C. Conversely, the estimation of the crowd density for a high number of people is higher in the biggest room rather than in the smallest one. This is because few people in a small environment move more closer to the transmitter and/or the receiver, so they cause a higher variation on the channel frequency response, making it distance feature more discriminant. Conversely, when a high crowd of people is in the environment, the moving area is spatially limited, thus distance tends to saturate its range loosing the ability to discriminate with high accuracy the crowd size.

	Prediction				
Truth	Empty	1 person	2 people	3,4 people	5,6,7 people
Empty	1	0	0	0	0
1 person	0	0.79	0.22	0	0
2 people	0	0.10	0.85	0.05	0
3,4 people	0	0	0.40	0.55	0.05
5,6,7 people	0	0	0.12	0.36	0.52
					0.74

Table 1: Confusion matrix for the Room A

	Prediction				
Truth	Empty	1 person	2 people	3,4 people	5,6,7 people
Empty	1	0	0	0	0
1 person	0	0.44	0.11	0.31	0.14
2 people	0	0.18	0.14	0.41	0.27
3,4 people	0	0.01	0.04	0.12	0.83
5,6,7 people	0	0	0	0.10	0.90
					0.52

Table 2: Confusion matrix for the Room C

5. RELATED WORKS AND ESTIMATION ACCURACIES

This Section provides an overview of the most important related works. They can be divided into two main categories: RSSI-based and CSI-based. As a matter of fact, most of the works are based on RSSI measurements.

Depatla et al. developed a mathematical expression for the probability distribution of the received signal amplitude as a function of the total number of occupants, which is the base for people counting using Kullback-Leibler divergence [3]. A single AP and a single receiver are used and the estimation delay is 300 s. Two sites are used in the experiments: an indoor site with the dimension 4.4×7.5 m^2 and an outdoor site with the dimension 7×10 m^2. Experiments performed with up to 9 people show that an error of 2 or less is achieved 96% and 63% of the time for the outdoor and indoor cases respectively, when using the typical omnidirectional antennas that come as part of the standard WiFi cards.

Xu et al. deployed an infrastructure of 20 to 22 devices to count and localize subjects in large areas [4]. The radio devices used in the experiments contain a Chipcon CC1100 radio transceiver operating in the unlicensed band at 909.1 MHz. Each transmitter periodically broadcasts a 10-byte packet every 100 ms. When the receiver receives a packet, it measures the RSS values and wraps the transmitter ID, receiver ID, RSS, timestamp (on the receiver side) into a service packet which is sent to a centralized system for data collection and analysis. Experimental results show that SCPL works well in two different typical indoor environments of 150 m^2 (office cubicles) and 400 m^2 (open floor plan). In both spaces, we can achieve about an 86% accuracy for up to 4 subjects.

Yuan et al. deployed a grid of 16 sensor nodes (TelosB) in a room with dimension 18×18 m^2 [5]. A database server collects the signal strengths and uses the information to cluster the crowd density level in the room. An average accuracy of 86% was achieved to estimate three types of densities are estimated: low (0-3 persons), moderate (4-10 persons) and high (>10 persons).

Fadhlullah et al. deployed three ZigBee wireless nodes continuously reporting the received signal strength indicator to the main node [6]. The average RSSI was calculated using the recorded information of all tags and classified into three crowd density categories which are low (0-5 persons), medium (6-10 persons), and high (11-15 persons). The results showed that the system was 75% and 70% accurate in detecting the low and medium human crowd density, respectively.

Yoshida et al. applied a regression-based approach to count the number of people using the RSSI collected by 10 nodes using one AP in a single room [7]. The estimation delay is 300 s. The accuracy of support vector regression-based method for estimating the number of people is 77% for up to 7 people.

The only CSI-based crowd counting method proposed in literature is [8]. They applied a Gray Verhulst model to construct a normal profile of the percentage of nonzero elements of the dilated matrix of CSIs for each class. The experiments are carried out counting up to 30 persons in both indoor and outdoor settings using a single AP and 3 or 4 receivers. The results show that this approach achieves an error ranging from 2 persons to three persons; more than 98% estimation errors are less than 2 persons in indoor environment, while about 70% errors is 2 persons in outdoor environment.

Comparison of performance and experimental setup among crowd counting methods including our method is shown in Table 3. It is worth noting that all the reported methods, except our method, perform a training phase in the same room where they perform the crowd counting test.

Table 3: Comparison of device-free crowd counting methods

Ref	Environments	Max # of People	Standards	Source Data	# TXs	# RXs	Estimation Delay (window size W)	Crowd Counting Accuracy	Crowd Density Accuracy
[3]	2 (indoor 33 m^2, outdoor 70 m^2)	9	WiFi	RSSI	1	1	300 s	P(e≤ 2) = 96%(outdoor) P(e≤ 2) = 63%(indoor)	-
[4]	2 (indoor 150 m^2, indoor 400 m^2)	4	ISM 909.1 MHz	RSSI	12/13	8/9	1 s	84%	-
[5]	1 (indoor 324 m^2)	>10	ISM 2.4 GHz	RSSI	16	16	-	-	86% (0p-3p, 4p-10p, >10p)
[6]	1 (indoor 100 m^2)	15	ZigBee 2.4 GHz	RSSI	1	3	-	-	73% (5p, 10p, 15p)
[7]	1 (indoor)	7	WiFi	RSSI	1	10	300 s	77%	94% (0p, 1p-3p, 3p-7p)
[8]	2 (indoor, outdoor)	30	WiFi	CSI	1	3/4	-	P(e≤ 2) = 70% (outdoor) P(e≤ 2) = 98% (indoor)	-
Our system	3 (indoor 30 m^2, indoor 45 m^2, indoor 70 m^2)	7	WiFi	CSI	1	1	10 s	P(e≤2) = 91% (room A) P(e≤2) = 81%(room C)	74% (room A) 52% (room C) (0p, 1p, 2p, 3p-4p, 5p-7p)

6. CONCLUSION

In conclusion, to the best of our knowledge, this is the first WiFi-based device-free crowd counting method which does not require to train the system exactly in the same room/environment where the system is also tested. First of all, the presented confusion matrices prove the effectiveness of our approach for the crowd counting and crowd density estimation. In fact, concerning crowd counting, estimation errors are less or equal to 2 persons 91% and 81% of time, while concerning crowd density estimation, average accuracies are 74% and 52% in the rooms where a specific training is not carried out.Some future directions of this work includes: study of the occupancy dynamics; estimation of larger crowd in larger rooms; evaluate the impact of the transmitter and receiver locations on the system performance.

7. REFERENCES

[1] Y. Wang, J. Yang, Y. Chen, H. Liu, M. Gruteser, and R. P. Martin, "Tracking human queues using single-point signal monitoring," in *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '14. New York, NY, USA: ACM, 2014, pp. 42–54. [Online]. Available:

[2] M. De Sanctis, E. Cianca, S. Di Domenico, D. Provenziani, G. Bianchi, and M. Ruggieri, "Wibecam: Device free human activity recognition through wifi beacon-enabled camera," in *Proceedings of the 2nd Workshop on Workshop on Physical Analytics*, ser. WPA '15, 2015.

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

[4] C. Xu, B. Firner, R. S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, and N. An, "Scpl: Indoor device-free multi-subject counting and localization using radio signal strength," in *Proceedings of the 12th International Conference on Information Processing in Sensor Networks*, 2013, pp. 79–90.

[5] Y. Yuan, J. Zhao, C. Qiu, and W. Xi, "Estimating crowd density in an rf-based dynamic environment," *Sensors Journal, IEEE*, vol. 13, no. 10, pp. 3837–3845, Oct 2013.

[6] S. Y. Fadhlullah and W. Ismail, "A statistical approach in designing an rf-based human crowd density estimation system," *International Journal of Distributed Sensor Networks*, vol. 2016, pp. 8 351 017:1–8 351 017:9, 2016. [Online]. Available:

[7] T. Yoshida and Y. Taniguchi, "Estimating the number of people using existing wifi access point in indoor environment," in *Proceedings of the 6th European Conference of Computer Science (ECCS '15)*, 2015, pp. 46–53.

[8] W. Xi, J. Zhao, X. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using wifi," in *2014 IEEE Conference on Computer Communications, INFOCOM 2014, Toronto, Canada, April 27 - May 2, 2014*, 2014, pp. 361–369.

[9] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM CCR*, vol. 41, no. 1, p. 53, Jan. 2011.