PCA-Based inversion of WiFi signal for robust device-free indoor target detetion

PCA-Based inversion of WiFi signal for robust device-free indoor target detection

Giorgio Gottardi [1], **Mohammad Abdul Hannan** [1], **Baozhu Li** [1,2], **Alessandro Polo** [1], **Marco Salucci** [1], **and Federico Viani** [1]

[1] ELEDIA Research Center (ELEDIA@UniTN - University of Trento), Via Sommarive 9, I-38123 Trento, Italy

[2] School of Physics and Technology, Nanjing Normal University, Wenyuan Road No.1, 210023 Nanjing, China

E-mail: giorgio.gottardi@unitn.it

Abstract. The spatial and frequency diversity introduced by the recent WiFi standards based on orthogonal frequency-division multiplexing is exploited for the robust wireless detection of device-free targets in indoor environments. The arising inverse problem is solved by decomposing the channel state information by means of a customized strategy based on the principal component analysis (PCA) for extracting target-dependent components from the WiFi signal. The experimental validation of the proposed solution pointed out a robust target detection with failure rates lower than 3 [%] using a single wireless link already deployed in a real office test field.

1. Introduction

The device-free wireless localization, also known as passive wireless localization, has attracted an increasing interest in the past few years [1]-[5]. A variety of location-based applications benefit from the detection and localization of transceiver-free (i.e., not carrying any device) targets, especially in complex indoor environments occupied by non-cooperative users. Many wireless technologies have been exploited with different levels of hardware customizations to acquire useful target-dependent features of the electromagnetic (EM) indoor propagation. The wireless sensor network (WSN) and the wireless fidelity (WiFi) technologies are two representative examples widely investigated in the state of the art [6]-[9]. Both of them provide the well-known received signal strength indicator (RSSI), which has been deeply analysed and exploited for passive localization thanks to its sensitivity to the perturbations caused by the environment as well as by the targets' presence and movements [10]-[12]. More recently, with the introduction of new WiFi standards like the IEEE 802.11n, which is based on the orthogonal frequency-division multiplexing (OFDM) and on multiple-input multiple-output (MIMO) antenna configurations, the channel state information (CSI) has been leveraged by detection and localization algorithms as a powerful indicator of the EM propagation at multiple subcarrier frequencies [13][14]. The CSI provides considerably higher information content than the RSSI since the frequency diversity and the spatial diversity are considered in the estimation of the "channel quality" between transmitters and receivers. The magnitude and the phase of the CSI have been exploited in the recent state of the art for wireless localization purposes [15]-[18]. However, most of the existing solutions aggregate the CSI features in order to define numerical indicators that are

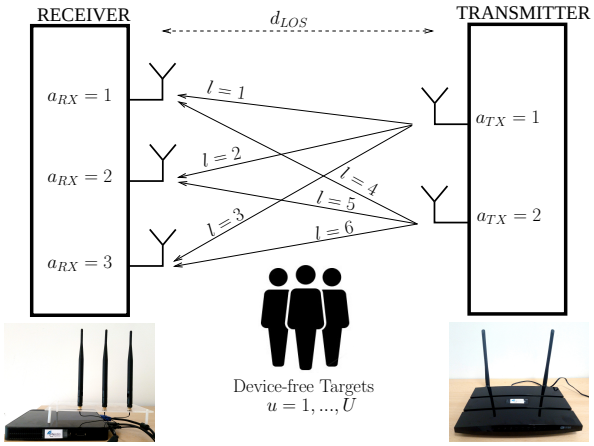

Figure 1. Architecture of the MIMO wireless system for device-free target detection.

somehow correlated with the presence and position of the targets. Less attention has been given to the exploitation of both the frequency and the spatial diversity in order to extract and isolate only the target-dependent features of the CSI and make the target detection more reliable. In this work, a strategy based on the processing of the CSI is proposed for the truly robust and almost *errorless* detection of passive targets. The method is aimed at separating the effects on the CSI caused by the target from the perturbations caused by the environmental noise and by the indoor propagation phenomena. Such a challenging goal is addressed by applying the principal component analysis (PCA) [19] on the outputs of the non-parametric Kruskal-Wallis (KW) test, which is adopted to estimate the similarity level between the CSI distributions acquired in absence and in presence of targets. To the best of the authors' knowledge, such an approach for the analysis of CSI data has never been proposed in the state of the art. Preliminary experimental results have pointed out the feasibility to separate those principal components of the CSI, which are target-dependent and mathematically orthogonal to the remaining signal components. The isolation of such a subset of principal components has led to a device-free target detection highly robust to environmental noise, even if a single wireless link between a couple of commercial WiFi transmitter-receiver is used. A set of experiments have been performed in a real office environment to detect the presence of users during daily activities. The obtained results pointed out outstanding detection performance with a failure rate lower than 3 [%]. The mathematical formulation of the proposed PCA-based approach is reported in Sect. 2, whereas Sect. 3 describes the experimental validation performed in a real indoor environment. Final conclusions and future activities on the development of the proposed solution are reported in Sect. 4.

2. Mathematical Formulation

Let us consider a WiFi transmitter and a receiver located in know positions $\underline{r}_{tx}$ and $\underline{r}_{rx}$, respectively, with the Euclidean norm $d_{LOS} = \|\underline{r}_{tx} - \underline{r}_{rx}\|$ being the line-of-sight (LOS) link length and $\underline{r} = (x, y, z)$ the position vector (Fig. 1). The transmitter is equipped with $a_{tx} = 1, ..., A_{tx}$ antennas, while $a_{rx} = 1, ..., A_{rx}$ receiving antennas are installed on the receiver. Such a MIMO configuration enables a total number of $L = A_{tx} \times A_{rx}$ wireless links. The radio-frequency (RF) propagation of each link is characterized by the CSI formulated as follows

$$h_l^{(c)}(t) = \alpha_l^{(c)}(t)\, e^{j \sin \varphi_l^{(c)}(t)}; \quad l = 1, ..., L;\ c = 1, ..., C \tag{1}$$

where $\alpha_l^{(c)}$ and $\varphi_l^{(c)}$ are the amplitude and the phase of the complex CSI values measured at the c-th frequency carrier of the l-th wireless link, and t is the acquisition time instant. The CSI amplitudes acquired within a time window are stored in the vector $\underline{\alpha}_l^{(c)} = \left[\alpha_l^{(c)}(t - w\Delta t)\right]$, $w = 1, ..., W-1$, where Δt is the sampling time interval and W the number of the last CSI samples collected in time and stored in the time window. The CSI samples are acquired both in *absence* (i.e., $\underline{\alpha}_l^{(c)} = \underline{A}_{l,c}$) and in *presence* (i.e., $\underline{\alpha}_l^{(c)} = \underline{P}_{l,c}$) of targets. The temporal correlations $\underline{\chi}_{l,c}^{A} = corr_w\left[\underline{R}_{l,c}, \underline{A}_{l,c}\right]$ and $\underline{\chi}_{l,c}^{P} = corr_w\left[\underline{R}_{l,c}, \underline{P}_{l,c}\right]$ are computed, where $\underline{R}_{l,c}$ is a reference absence acquisition, and $corr_w\left(\underline{x}, \underline{y}\right) = \frac{cov(\underline{x},\underline{y})}{\sigma_{\underline{x}}\sigma_{\underline{y}}}$, where $cov\left(\underline{x}, \underline{y}\right)$ is the covariance and σ the standard deviation. The corresponding probability density functions (PDFs) $p\left(\underline{\chi}_{l,c}^{A}\right) = \sum_{w<W} \underline{\chi}_{l,c,w}^{A}$ and $p\left(\underline{\chi}_{l,c}^{A}\right) = \sum_{w<W} \underline{\chi}_{l,c,w}^{A}$ are computed to extract the patterns of the CSI amplitudes in different *absence*/*presence* status. It has been verified in [16] that such PDFs are temporally stable in static environments, whereas perturbations arise in presence of target motions. In the state of the art, the target-dependent behaviour of CSI has been often exploited merging and averaging the information available in the frequency domain (i.e., over the $c = 1, ..., C$ frequency carriers) and in the spatial domain (i.e., considering the $l = 1, ..., L$ MIMO links), but without clearly isolating the target-dependent features of the CSI. In order to address this challenge, the non-parametric KW test has been adopted. The KW test is commonly adopted in statistics to test whether two or more samples originate from the same distribution without the assumption that the residuals have a normal distribution [17]. The proposed approach applies the KW test to determine the similarity level between the non-normally distributed PDFs $p\left(\underline{\chi}_{l,c}^{A}\right)$ and $p\left(\underline{\chi}_{l,c}^{P}\right)$, $l = 1, ..., L$, $c = 1, ..., C$, computed for all the frequency carriers and for all the wireless links. More in detail, the p-values $\rho_{l,c}$, $l = 1, ..., L$, $c = 1, ..., C$, given by the KW test applied to the whole set of $C \times L$ PDFs have been computed at each sampling time instant. If $\rho < \varepsilon_{KW}$, ε_{KW} being the significance level threshold, the difference between the PDFs is statistically significant, otherwise they belong to the same family of distributions. In order to identify and extract the most target-dependent features from the computed p-values, the PCA has been applied to the set of p-values ρ_l, $l = 1, ..., L$, and the procedure has been repeated for each carrier frequency $c = 1, ..., C$.

The PCA is defined as an orthogonal linear transformation of data into a new coordinate system such that the greatest variance of data projection lies on the new coordinates in decreasing order of magnitude. The transformation is defined by a set of weight vectors (also called PCA *coefficients*) mapping each data vector to a new vector of principal vector scores $\underline{\tau}_c = [\tau_l;\, l = 1, ..., L]_c$. Such an orthogonal transformation is aimed to extract the linearly uncorrelated components existing among the wireless links $l = 1, ..., L$, pointing out which subset of components is the most affected by the target presence, across the considered frequency spectrum. More in detail, the principal components scores

$$\underline{\tau}_c = [\tau_l;\, l = 1, ..., L]_c = PCA\left\{\rho_{l,c};\, l = 1, ..., L\right\}; \quad c = 1, ..., C \tag{2}$$

are computed by the method, where $PCA\{\cdot\}$ is the PCA-based transformation of input p-values providing in output the scores of the principal components [19]. The obtained components $\underline{\tau}_c$, $c = 1, ..., C$, are representative indicators of the unknown spatial relations existing among the multiple wireless links at the different carrier frequencies. Low values of τ_l, $l = 1, ..., L$ are an indication of the target presence, which is the cause of lower spatial correlation among the PDFs of the processed links. Consequently, the temporal analysis of the τ_l patterns enables the real-time target detection.

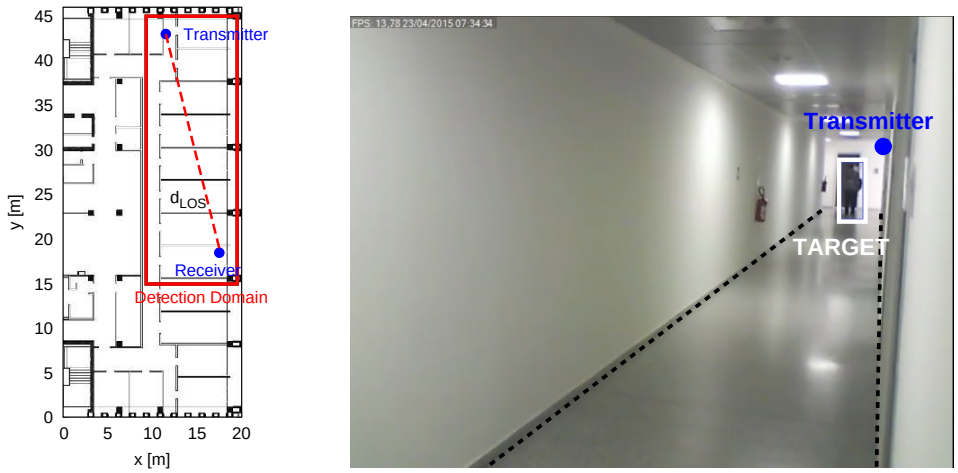

Figure 2. (a) Indoor domain for experimental validation of the target detection using one wireless link, and (b) example of ground truth acquired for detection performance assessment.

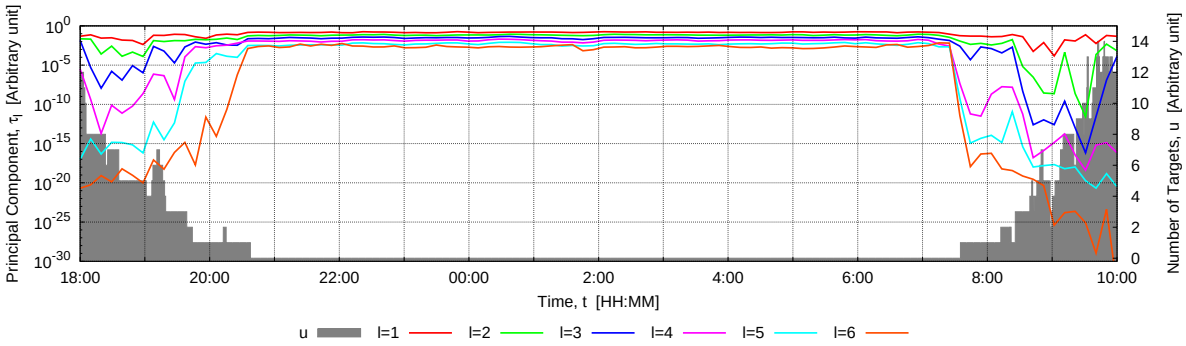

Figure 3. Temporal evolution of the principal components τ_l; $l = 1, ..., L|_{c=11}$, $L = 6$, versus the time-varying target *absence*/*presence* ($u = 1, ..., U$, $U = 13$).

3. Experimental Validation

The proposed method for device-free target detection has been experimentally validated in a real indoor test field at the ELEDIA Research Center laboratories, University of Trento, Italy. The office area reported in Fig. $2(a)$ has been selected for the target detection experiments. A commercial WiFi access point (AP) compliant to the IEEE 802.11n standard and with $A_{tx} = 2$ antennas has been used as transmitter, while the receiver has been equipped with the Intel IWL 5300 network interface card (NIC) [20] using $A_{rx} = 3$ antennas. The devices have been installed at the ceiling level ($z = 2.6$ $[m]$). A total of $L = 6$ wireless links working at the center frequency $f_c = 2.4$ $[GHz]$ and LOS distance $d_{LOS} = 17$ $[m]$ between positions $\underline{r}_{tx}$ and $\underline{r}_{rx}$ have been established in a non-line-of-sight (NLOS) scenario. Thanks to the OFDM adopted by the wireless communication standard and the features of the receiving NIC, $C = 30$ frequency carriers have been measured during the experiments. The CSI amplitudes have been acquired with a sampling rate $\Delta t = 0.5$ $[s]$ and stored in a sliding window of $W = 40$ samples. The size of the time window has been calibrated to maximize the statistical value of the PDFs as well as to ensure a short time delay for real-time detection performance. The reference *absence* acquisition

Table 1. Detection performance versus the principal components τ_l $l = 1, ..., L|_{c=11}$, $L = 6$.

Principal Components τ_l	False Negative [%]	False Positive [%]	Failure Rate δ [%]
$l = 1$	99.6	0.0	50.0
$l = 2$	81.4	0.0	40.7
$l = 3$	64.2	0.0	32.1
$l = 4$	33.6	0.0	16.8
$l = 5$	21.4	0.4	10.9
$l = 6$	3.8	1.2	2.5

has been stored during the night-time with empty indoor site, while the remaining acquisitions have been acquired with different target *absence/presence* status, lasted 16 hours during night and day times. During this time period, the considered domain has been occupied by a variable number of $u = 0, ..., U$ targets doing regular office activities. The actual target number has been verified by means of a video-surveillance system recording the detection area as shown in Fig. 2(b). The KW test has been performed on the $C \times L = 180$ PDFs and it has been iterated every Δt sampling interval for the whole duration of the experiment. The obtained p-values have been compared with the significance threshold $\varepsilon_{KW} = 0.05$ for preliminary filtering, and successively they have been adopted as input variables of the PCA. The principal components $[\tau_l; l = 1, ..., L]$ have been computed for all the carriers $c = 1, ..., C$, $C = 30$. For the sake of brevity, one representative carrier ($c = 11$) has been selected among the others in order to present the results of the detection. The time evolution of the principal components τ_l, $l = 1, ..., L$, $L = 6$, is reported in Fig. 3 together with the actual number of targets $u = 0, ..., U$, $U = 13$. As it can be noticed, an evident relation between the target *absence/presence* and the principal components index has been obtained. To the best of the author's knowledge, the relation between the target presence and the behaviour of such orthogonal components based on the spatial diversity of the MIMO wireless links has not been investigated in the state of the art. The first components (e.g., $l = 1$ and $l = 2$) are insensitive to the target presence, while the last ones (in particular the component $l = 6$), are highly target-dependent even when the monitored area is occupied by a single target (i.e., $u = 1$). According to the PCA theory, the first components are those with the largest possible variance of the input data. Since the values of τ_l, $1 \leq l \leq 3$, are almost stable regardless the target *absence/presence*, the input data variability related to the first subset of components has been associated to the noise of the complex indoor scenario, whereas the target *absence/presence* has been inferred from the behaviour of the smallest components (e.g., $4 \leq l \leq 6$). Accordingly, the robustness of the detection has been improved filtering the noisy components and processing only the target-dependent ones.
The detection performance have been analysed in terms of failure rate δ [%], which has been computed as the average value of false positive and false negative detections. The analysis has been performed for each principal component assuming a simple thresholding strategy, with the user-defined threshold $\tau_{th} = 10^{-5}$ (i.e., the target is detected if $\tau < \tau_{th}$). The outstanding detection performance of the target-dependent principal component τ_6 respect to the preceding ones is clearly pointed out by the results reported in Tab. 1. The failure rate rapidly decreased down to $\delta = 10.9$ [%] with $l = 5$, and to $\delta = 2.5$ [%] with $l = 6$.

4. Conclusions

The detection of device-free targets has been addressed by means of an innovative inversion computational method able to identify and extract the target-dependent features of the CSI.

The spatial diversity of a MIMO IEEE 802.11n wireless link has been exploited to isolate the orthogonal signal components related to the target presence. The experimental validation has pointed out a failure rate lower than 3 [%] when the noisy components of the CSI are filtered out. It has to be noticed that such a robust detection has been obtained by simply applying a threshold to the principal component values. The proposed method has pointed out the feasibility to integrate a robust target detection feature on top of existing WiFi architectures already deployed for indoor wireless connectivity. A more detailed analysis of the frequency diversity is currently under study to further exploit the properties of the whole CSI frequency spectrum toward an almost errorless detection of passive targets.

Acknowledgments

This work benefited from the networking activities carried out within the Project "CYBER-PHYSICAL ELECTROMAGNETIC VISION: Context-Aware Electromagnetic Sensing and Smart Reaction (EMvisioning)" funded by the Italian Ministry of Education, University, and Research within the PRIN2017 Program, and the Project "SMARTOUR - Piattaforma Intelligente per il Turismo" (Grant no. SCN_00166) funded by the Italian Ministry of Education, University, and Research within the Program "Smart cities and communities and Social Innovation".

References

[1] Wilson J and Patwari N 2012 A fade-level skew-laplace signal strength model for device-free localization with wireless networks *IEEE Trans. Mobile Comput.* **11** 947-958
[2] Viani F, Migliore M D, Polo A, Salucci M and Massa A 2018 Iterative classification strategy for multi-resolution wireless sensing of passive targets *Electron. Lett.* **54** 101-103
[3] Li Q, Fan H, Sun W, Li J, Chen L and Liu Z 2017 Fingerprints in the air: unique identification of wireless devices using RF RSS fingerprints *IEEE Sensors J.* **17** 3568-3679
[4] Viani F, Robol F, Polo A, Rocca P, Oliveri G and Massa A 2013 Wireless architectures for heterogeneous sensing in smart home applications - Concepts and real implementations *Proc. IEEE* **101** 2381-2396
[5] Tayebi A, Gomez Perez J, Adana Herrero F M S and Gutierrez O 2009 The application of ray-tracing to mobile localization using the direction of arrival and received signal strength in multipath indoor environments *Prog. Electromagn. Res.* **91** 1-15
[6] Viani F, Polo A, Garofalo P, Anselmi N, Salucci M and Giarola E 2017 Evolutionary optimization applied to wireless smart lighting in energy-efficient museums *IEEE Sensors J.* **17** 1213-1214
[7] Viani F, Bertolli M and Polo A 2017 Low-cost wireless system for agrochemical dosage reduction in precision farming *IEEE Sensors J.* **17** 5-6
[8] Ahmadi H, Polo A, Moriyama T, Salucci M and Viani F 2016 Semantic wireless localization of WiFi terminals in smart buildings *Radio Sci.* **51** 876-892
[9] Viani F, Polo A, Donelli M and Giarola E 2016 A relocable and resilient distributed measurement system for electromagnetic exposure assessment *IEEE Sensors J.* **16** 4595-4604
[10] Mrazovac B, Todorovic B M, Bjelica M Z and Kukolj D 2013 Device-free indoor human presence detection method based on the information entropy of RSSI variations *Electron. Lett.* **49** 1386-1388
[11] Xue W, Qiu W, Hua X and Yu K 2017 Improved Wi-Fi RSSI measurement for indoor localization *IEEE Sensors J.* **17** 2224-2230
[12] Wu C, Yang Z, Zhou Z, Liu X, Liu Y and Cao J 2015 Non-invasive detection of moving and stationary human with wifi *IEEE J. Sel. Areas Commun* **33** 2329-2342
[13] Chapre Y, Ignjatovic A, Seneviratne A and Jha S 2015 CSI-MIMO: an efficient Wi-Fi fingerprinting using channel state information with MIMO *Pervasive Mob. Comput.* **23** 89-103
[14] Fang S, Chang W, Tsao Y, Shih H and Wang C 2016 Channel state reconstruction using multilevel discrete wavelet transform for improved fingerprinting-based indoor localization *IEEE Sensors J.* **16** 7784-7791
[15] Wu K, Xiao J, Yi Y, Chen D, Luo X and Ni L 2013 CSI-based indoor localization *IEEE Trans. Parallel Distrib. Syst.* **24** 1300-1309
[16] Xiao J, Wu K, Yi Y, Wang L and Ni L M 2013 Pilot: Passive device-free indoor localization using channel state information *IEEE Int. Conf. Distributed Computing Systems*, Philadelphia, PA, USA, 2013, 236-245
[17] Wayne W D 1990 Kruskal-Wallis one-way analysis of variance by ranks *Applied Nonparametric Statistics (2nd ed.)* Boston 226-234

[18] Wang X, Gao L, Mao S and Pandey S 2017 CSI-based fingerprinting for indoor localization: a deep learning approach *IEEE Trans. Veh. Technol.* **66** 763-776
[19] Fang S and Lin T 2012 Principal component localization in indoor WLAN environments *IEEE Trans. Mobile Comput.* **11** 100-110
[20] Halperin D, Hu W, Sheth A and Wetherall D 2010 Predictable 802.11 packet delivery from wireless channel measurements *Proc. ACM SIGCOMM*, NY, USA, 2010, 159-170