Passive Device-free Indoor Movement Detection using WiFi Time-of-Flight Information

Sascha Rösler, Anatolij Zubow and Falko Dressler
School of Electrical Engineering and Computer Science, TU Berlin, Germany
{roesler, zubow, dressler}@tkn.tu-berlin.de

***Abstract*—We present ToFMD, a system that can passively detect the movement of an indoor device-free user by relying on the absorptions and reflections of WiFi signals off his body. Our system is of practical use as it is able to reuse existing deployed commodity 802.11 infrastructure networks. It is based on the observation that the movement of an user is altering the multipath characteristics of the wireless channel which leads to changes in the measured time-of-flight (ToF) between adjacent WiFi access points. The system can detect moving people at different places in the room, it is not necessary that the user blocks the line-of-sight (LOS) path. In contrast to other approaches, our system works with inexpensive single antenna WiFi systems without access to low-level parameters like channel state information (CSI). Instead the ToF between any pair of access points is estimated using the fine time measurement (FTM) protocol, which is standardized within IEEE 802.11. The movement is detected with the help of a one class support vector machine. As proof of concept, we implemented our system with ultra-low cost ESP32 microcontrollers. The evaluation shows that our solution outperforms traditional approaches using RSSI except for the scenario with user motion in adjacent rooms. If our system is installed in the same room as the user, it is able to detect its movement with an accuracy of at least 95 %.**

***Index terms*—** wireless sensing, indoor movement detection, device-free detection, time-of-flight, IEEE 802.11 FTM

I. Introduction

There is rich literature body on device-free presence detection schemes based on IEEE 802.11 WiFi [1]. However, they either rely on complex hardware (MIMO arrays) and/or special software. Some approaches rely on processing of channel state information (CSI) which is available with COTS WiFi hardware, however, require either customized device drivers or root access to the WiFi devices. Other approaches rely on processing of basic parameters like received signal strength indication (RSSI) which are available to normal applications, e.g., in Android operating system. However, these approaches suffer from poor performance in certain environments [2].

Results from experiments reveal that our `ToFMD` outperforms traditional RSSI-based approaches in most scenarios except for user movement in adjacent rooms, i.e., behind a wall. In this paper, we present `ToFMD`, a passive approach for device-free movement detection of users indoors by processing time-of-flight (ToF) data between any pair of adjacent and cooperating WiFi access points (APs) (Fig. 1). As ToF-based ranging is standardized within IEEE 802.11, information on ToF is available to normal applications like RSSI. Our approach is based on the observation that the movement of an user is altering the multipath characteristics of the wireless channel which lead to changes in ToF measurements. For this it is not necessary that the user blocks the line-of-sight (LOS) path. Moving behind or next to the LOS link also changes the measured ToF. In contrast to other approaches our system works with commodity single antenna hardware without access to low-level parameters like CSI, i.e., unrooted devices. As proof of concept, we implemented our system with ultra-low cost ESP32 microcontrollers which are supporting 802.11n and the Fine Time Measurement (FTM) standard required for ToF-based ranging. Even though the microcontrollers also provide CSI measurements, we are not using it as we want to provide a system which is usable on Android devices and desktop computers without root access.

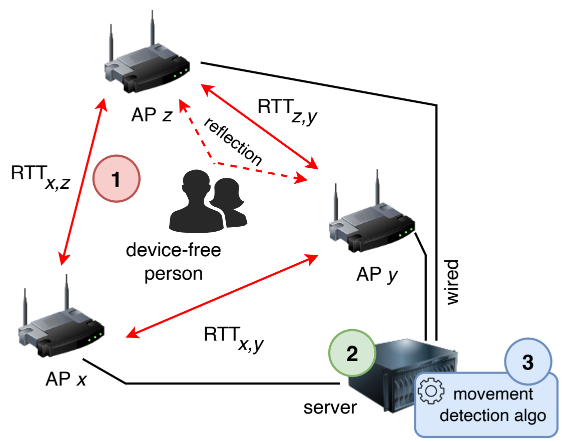


Fig. 1: `ToFMD` architecture: 1) ToF estimation between fixed installed WiFi APs using 802.11 FTM protocol, 2) data fusion at server, and 3) algorithm for device-free movement detection.

Our approach is of ultra-low complexity and low overhead. The ToF ranging between WiFi APs is performed at very low rate, i.e., every 2s.

Our contributions can be summarized as follows:

- Experimental results reveal the influence of a human body on the measured ToF between a pair of WiFi devices that can be exploited for movement detection.
- We propose `ToFMD`, a practical approach for device-free indoor movement detection utilizing one class support vector machine (OC-SVM).
- Our implementation with ultra-low cost COTS WiFi hardware outperforms traditional RSSI-based approaches in multiple scenarios.

II. Related Work

Related work falls into four categories:

RSSI-based Designs: RSSI is used by WiDet [3] and by the approach of Depatla et al. [2]. WiDet achieves an accuracy of 94 % in detecting pedestrians with the help of a convolutional neural network and continuous wavelet transformation to detect slow moving persons [3]. Depatla et al. [2] use the RSSI to estimate the average speed of a crowd based on cross correlation and a Markov chain model. They build their prototype with commodity WiFi hardware.

CSI-based Designs: Natarajan et al. [4] use both, CSI and RSSI, to achieve movement detection using ESP32 WiFi hardware. While they use the RSSI value as a feature as itself the CSI data is filtered and features like variance and mean value are extracted. Using different learning models, e.g., random forest gradient boosting, they achieve an average detection accuracy based on RSSI of 83 %. Wu et al. [5] use the CSI values to calculate Doppler frequency shifts. Their approach works with Intel WiFi cards and achieves an accuracy of 88 %. TTW [6] also uses the calculation of Doppler frequency shifts to achieve an accuracy in movement detection of 99%. The same hardware is used in DeMan [7] which can detect moving humans with an accuracy of 95%. This approach uses Eigenvalues of the covariance matrix of consecutive CSI measurements. The authors of Wi-PSG [8] feed time and frequency features together into a multiclass SVM to detect rhythmic movements with an accuracy of 98 %. The CSI data was obtained from Intel WiFi cards, where the CSI was filtered with the help of a Hampel filter. Wang [9] use the CSI values for a breath rate detection. They also show the limits of CSI-based sensing based on a study of Fresnel zones. RT-Fall [10] is a fall detection solution that can detect different types of movement (like sitting, walking, standing) with an accuracy between 80 % and 90 %. It uses the difference in the phase information and calculates features like standard deviation and signal strength out of it to use them in a ν-SVM with an RBF kernel. FullBreath [11] uses the amplitude and the phase data of the CSI to measure the breath frequency. Only the magnitude of the CSI data is used by Korany et al. [12] to identify multiple persons moving around in a room with an accuracy of 82%. They calculate the angle of arrival and analyze the data by a short term Fourier transformation. In [13] the same authors count a crowd with an accuracy of 96%. They use the analogy to a $M/G/\infty$ queuing theory problem and estimate the number of persons based on busy (movement) times and idle times. Their solution bases on a Poisson distribution. WiSign [14] can recognize sign language with an accuracy of 87 %. A SVM or a k-nearest neighbors detector use features like mean and maximum value of the amplitude and the velocity of signal change.

MIMO-based Designs: Kianoush et al. [15] extract features like the mean or standard deviation, the skewness and the kurtosis from the data and estimate the number of persons based on a Kullback-Leibler divergence with an accuracy of 100 %. They also use a fast forward neuronal network and a long short term memory. Another solution based on CSI in MIMO systems is MoSense [16] that detects the movement of a person with an accuracy of 97 %. It also uses the difference in the phase information. In our previous work we presented an approach based on CSI, MIMO and the usage of OC-SVM working with normalized batches [17].

Specialized Waveforms: Some technologies use specialized waveforms for sensing and require specialized hardware like an USRP. In this way, Tan et al. [18] use a radar system to detect gestures based on Doppler frequency shifts with the help of a cross correlation function. WiZ [19] create a radar system that allows centimeter range localization estimation, gesture recognition and breath detection based on ToF data. It uses frequency modulated carrier waves in a MIMO system. With the help of silhouette canceling it can detect the position of different persons within 8 cm. The breath rate is correct in 97 % of the cases.

III. Background

A. Wireless Channel Propagation

A radio signal propagating through the wireless channel to the receiver over multiple paths experiences several effects. Reflection on walls or scattering through obstacles will produce additional copies of the transmitted signal, so-called multipath components, which have different arrival times. This behavior of the channel can be seen by different impulses in the channel impulse response (CIR). When assuming a time-invariant channel the CIR can be denoted as [20]:

$$h(t) = \sum_{n=0}^{N} a_n e^{-j\phi_n} \delta(t - t_n) \tag{1}$$

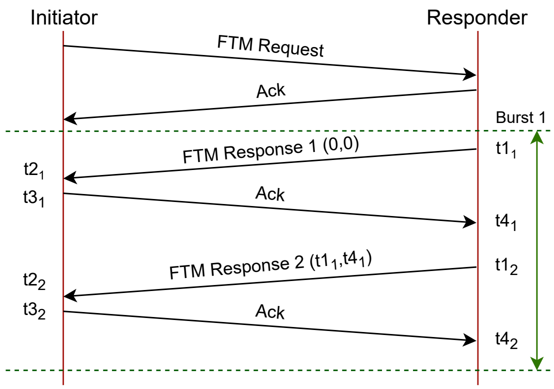


Fig. 2: An FTM exchange with one burst according to the IEEE 802.11 standard [21]

where a_n is the amplitude, ϕ_n is the phase of the nth multipath component at the time t, N is the total number of multipath components and $\delta(t)$ is the Dirac delta function. The sum of amplitude and phase of these multipath components can result in constructive or destructive interference.

B. Time-of-flight based Ranging with WiFi

The 802.11 fine time measurement (FTM) protocol [21] enables high accuracy ranging between any pair of WiFi stations. Synchronized clocks are not required as FTM uses a two-way time transfer protocol. In the protocol, one station acts as initiator while the other station is configured as responder. Association of the initiator with the responder is not required in order to perform ranging. This allows measurements to multiple responders.

High accuracy ranging in FTM is achieved by taking precise timestamps at the physical layer in picoseconds (ps) resolution, which gives them an accuracy of 0.03 cm. According to the standard [21], the timestamps should be taken as soon as the start of the preamble has been detected to make these timestamps as accurate as possible. The FTM protocol is defined as follows (Fig. 2): The responder takes the timestamps t_1 and t_4. It takes the first one immediately before it starts transmitting the FTM response frame while it takes the second when it receives the corresponding ACK frame. The initiator determines the timestamps t_2 and t_3, when the FTM response was received and when the initiator started transmitting the corresponding ACK frame respectively. These four timestamps form an FTM dialog for which the round trip time (RTT) can be calculated as:

$$\mathrm{RTT} = (t_4 - t_1) - (t_3 - t_2) \tag{2}$$

The estimated RTT is affected by fluctuations due to the limited bandwidth and environmental influences [22], [23]. Hence, in order to get higher accuracy, averaging is performed over multiple RTT values.

IV. Impact of Humans on ToF-based Ranging

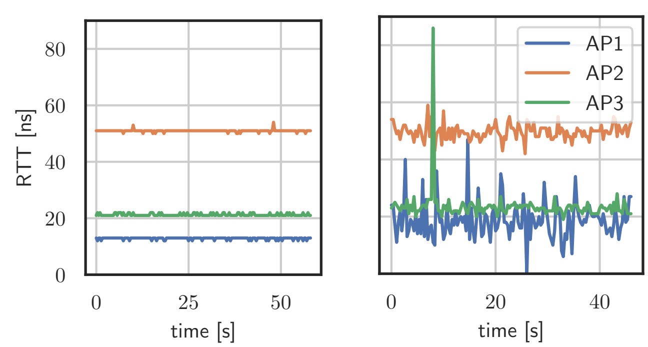
We start our study of how a human affects the RTT measurement of FTM with an over-the-air experiment in our lab. Therefore, we use the ESP32 SoC hardware platform supporting the 802.11 FTM protocol. The setup of the experiment shown in Fig. 9c consists of a single target and three anchor nodes. During the experiments the three anchor nodes performed ToF-based ranging towards the target node with a rate of 3 Hz. We analyzed the impact on the measured RTT for two scenarios: i) empty room and ii) a moving person inside that room.

The results are shown in Fig. 3. We can observe that in an empty room the measured RTT towards the three anchor nodes is stable fluctuating with only 1 ns which is the quantization error of the ESP32.

This changes dramatically when a person is moving inside the room resulting in sudden changes in measured RTT. Note, that even so the person was able to block at most the LOS path towards a single anchor we see fluctuations in RTT towards all three anchors.

Mathematically speaking, we see a clear difference when we calculate the standard deviation σ of RTT for both scenarios (see Table I). The σ for a moving person inside a room is one magnitude larger (4.84 ns) than it is for the empty room (0.39 ns). In addition we calculated σ for the perfect wired channel which is just 0.197 ns. This is even smaller and shows that the over-the-air transmission is impacted by additional external influences, e.g., micro-scale vibrations (windows) and co-channel interference.

For a deeper analysis Fig. 4 illustrates the three possible cases where a device-free human is affecting the wireless signal propagation of a point-to-point communication. First,



(a) empty room (b) moving person

Fig. 3: RTT in over-the-air measurements (lab, cf. Fig. 9c)

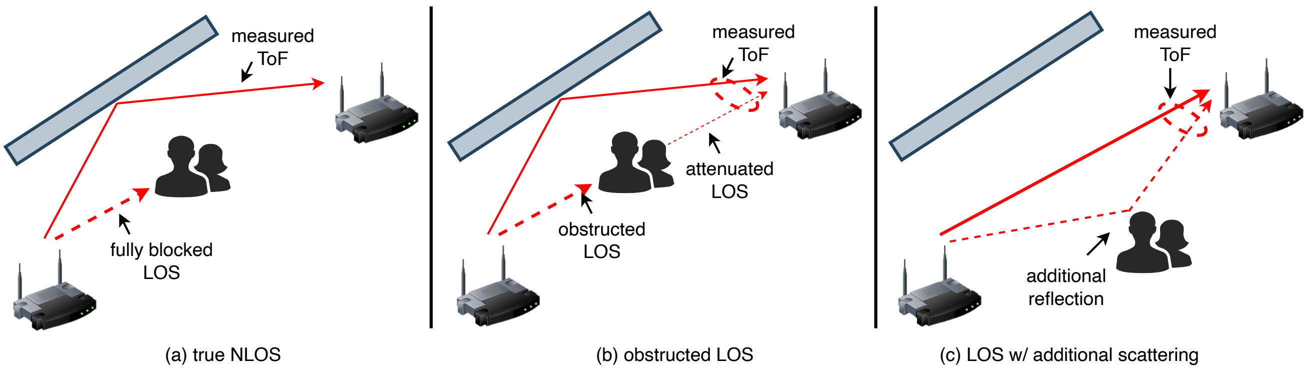


Fig. 4: Presence of a device-free human impacting round trip time (RTT) measurement.

by fully blocking the LOS path the signal is propagated over a true non-line-of sight (NLOS) channel. Here, the ToF of the reflection is measured instead of the shorter LOS leading to a significant alteration in measured ToF. Second, we show the case of obstructed, i.e., attenuated LOS path, in which the reflected paths might be stronger than the obstructed LOS one. Here, the ToF of the combined channel, i.e., weakened LOS with reflections, is measured which can lead to changes in measured RTT as well. Third, the human is not blocking the LOS path. Instead it creates an additional source of signal scattering resulting in a change of multipath propagation as we see in the results of our experiment. In this case, again, the ToF of the combined channel is measured.

Clearly, in the first case the human will influence the resulting measurement as the reflection is the strongest signal at the receiver. Hence, in the following we will study the impact for the last two cases. Therefore, we have already shown in our previous work [24] that ToF-based ranging of commodity 802.11 WiFi hardware is affected by multi-path propagation. Our assumption is, that the preamble detection algorithm is not capable to detect the start of a packet accurately when receiving multiple copies of the same packet.

A. Emulated (Wired) Multipath Channel

To get a better understanding of how the WiFi preamble detection behaves under different conditions, we conducted additional experiments using the ESP32 SoC hardware platform in a controlled wired setup. So we can study effects

TABLE I: Standard deviation of measured RTT in different scenarios.

scenario	σ [ns]
wired cable	0.197
empty room	0.389
moving target	4.839

TABLE II: Correlation between channel properties and RTT.

property	correlation coefficient	p-value
Multipath (d_2/d_1)	$r = 0.820$	$6.20 \cdot 10^{-4}$
Signal strength [dBm]	$r = 0.915$	$1.05 \cdot 10^{-2}$

caused by multipath and effects caused by signal attenuation separately.

First, we conducted measurements with ESP32 SoCs connected over an emulated 2-tap wired multipath channel (Fig. 5). Here we analyzed the impact of the ratio of the cable lengths, d_1 and d_2, on the measured RTT. From the results shown in Fig. 6 we can see that the increase in d_2/d_1 tends to increase the measured RTT. To assess the degree of correlation among the two metrics, we computed the Pearson correlation coefficient which was $r = 0.820$ (Table II).

In addition to multi-path, the packet detection accuracy is also impacted by the quality of the received signal which might also result in too early or too late packet detection. Therefore, we conducted additional experiments where we removed the first wired path (d_1) and analyzed the impact of attenuation on the remaining second path. The results are shown in Table III. With lower P_{rx} the RTT distribution becomes wider, i.e., larger σ, and the mean value μ is shifted to the right resulting in an overestimation of the true RTT. The Pearson correlation coefficient between signal strength (in dBm) and RTT was $r = 0.946$ (Table II).

In a third experiment we create a 2-tap wired multipath where one path (d_2) is attenuated (Fig. 5). We analyse the

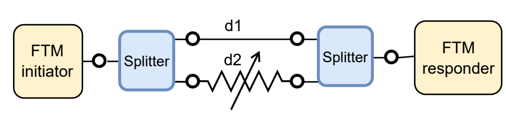


Fig. 5: Emulated 2-tap multipath channel over cable.

TABLE III: Impact of receive power P_{rx} on RTT.

P_{rx} [dBm]	μ	σ
-38	0.0	1053
-48	-575	1091
-58	280	1220
-68	4733	1634
-78	9809	3836
-88	17858	10657

effect on the measured RTT as well, as the measured RSSI. By comparing Figures 7a and 7b, we come to the conclusion that the RSSI is less affected by multipath than ToF (RTT), i.e., see orange curve.

V. ToFMD Approach

Our results from the previous section confirm that the RTT measured using ToF and commodity 802.11 hardware is highly affected by changes in the multipath channel propagation. This can be utilized to implement a passive device-free human movement detection scheme. Our key idea is to reuse existing WiFi infrastructure as much as possible. Therefore, we measure the RTT continuously between any pair of installed WiFi APs (Fig. 1). The measured RTT data streams from K APs are forwarded to a central server node where the movement detection algorithm is executed. Fig. 8 shows the processing pipeline.

A. In a Nutshell

The streams of RTT data from K-different APs (x) are pre-processed to extract the desired features. Here, two different processing chains (chain A and chain B) can be used. While the best results are achieved with chain A, chain B acts as baseline and gives best results when working with RSSI instead of RTT data. However, both processing chains can be used with both types of data (RTT or RSSI). Both chains are using a sliding window consisting of $B_t = 10$ data samples to create batches of data. Next chain A calculates the variance of a batch, while chain B normalizes the batch. The so extracted features are fed as flattened vector into the OC-SVM which performs the classification. Trained with the data of an empty room, the OC-SVM classifies whether the room is empty or if somebody is moving in the room, i.e., novelty.

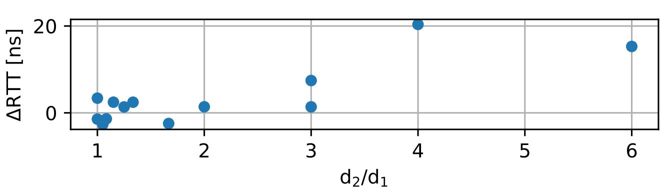
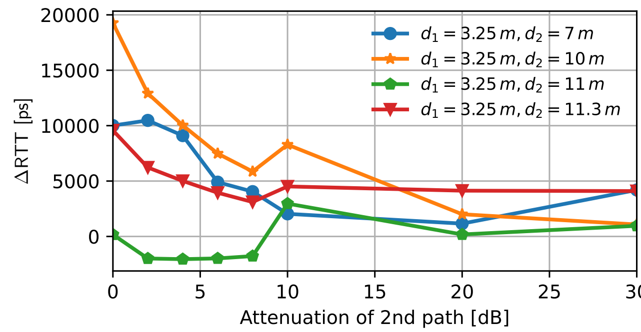
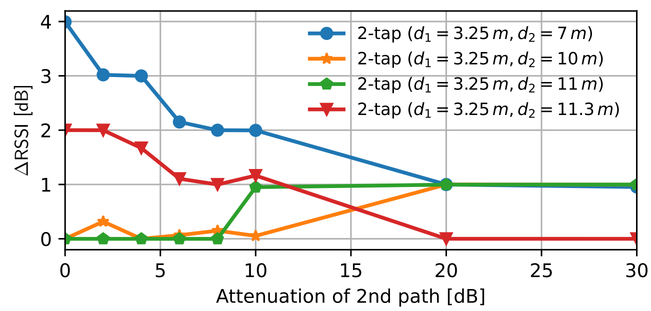


Fig. 6: Impact of d_2/d_1 on measured RTT from emulated multipath channel over cable.



(a) ToF (RTT)



(b) RSSI

Fig. 7: Impact of attenuation on second path on measured RTT/RSSI from emulated multipath channel over cable.

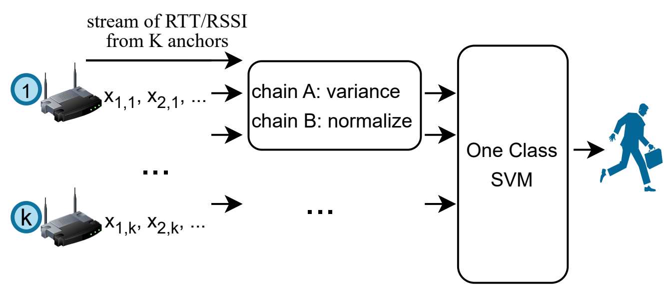
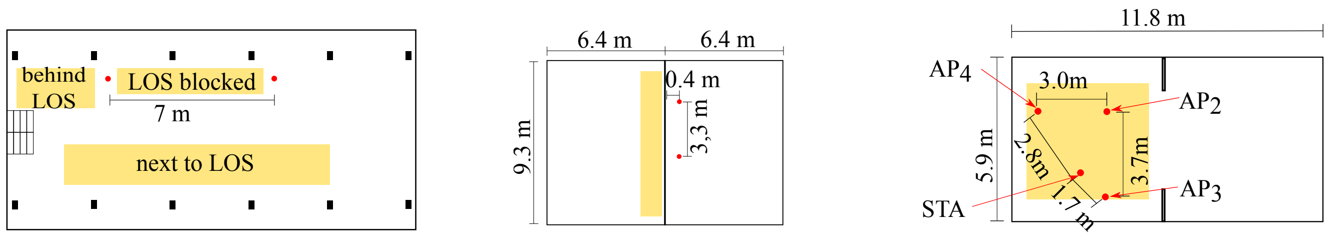


Fig. 8: `ToFMD` takes input from K anchor nodes and having two possible processing chains (chain A and chain B).

B. Detailed Description

Our processing pipeline consists of three major steps:

1. Pre-processing: The stream of RTT/RSSI data from the K anchor nodes is captured with a sample rate of $S = 3$ Hz. The data must be pre-processed before it can be passed to the OC-SVM for classification. First, we group our N RTT/RSSI samples into N_t batches of size B_t, with each batch resembling a time window of $\frac{B_t}{S}$. This results in the data dimension of



(a) hall (b) behind the wall (c) lab

Fig. 9: Indoor environments under study: a) hall with clear LOS condition between nodes (S1), b) adjacent room scenario (S2) and c) lab environment with multiple APs (S3). The location of tags and anchors is marked in red. The area where the person was moving is marked in yellow.

our data array X, given by

$$X = N_t \times B_t \times K \tag{3}$$

Within the duration of a batch, we expect a change in the temporal domain of RTT or RSSI due to a moving person. The next step is different for processing chains A and B. Without loss of generality we consider the i-th batch.

2.1 Feature extraction - Chain A: In mode A we simply compute the variance of each grouped batch B_t:

$$X_{i,:,:} \xrightarrow{\text{var()}} X^{\text{flat}} \tag{4}$$

This changes the dimension of the data into a vector of size K which serves as feature vector in next step.

2.2 Feature extraction - Chain B: In mode B we normalize each grouped batch B_t in respect to the first RTT/RSSI sample by element-wise division (Hadamard division). This gives us

$$X^{\text{norm}} = X_{i,:,:} \div X_{:,0,:} \tag{5}$$

Now the first entry of any batch B_t consists only of a vector with the value *1*, while the remaining entries of B_t have a relative value. This is done because the absolute RTT/RSSI value is dependent on the distance between transmitter and receiver and the multipath propagation conditions. With this normalization, we try to remove this environmental impact.

In addition we have to flatten our data dimensions of $B_t \times K$ into one dimension, which serves as feature vector for the input to the OC-SVM in the next step:

$$X^{\text{norm}} \xrightarrow{\text{flatten}} X^{\text{flat}} \tag{6}$$

3. Classification: The feature vector generated by either chain A or B is fed into the OC-SVM for classification. Here we consider data corresponding to no-mobility as regular and data corresponding to mobility as irregular, a novelty.

During training, the OC-SVM adjusts a function based on the training data. Therefore, it calculates bases in combination with a kernel specifying the shape of the function. E.g., the radial basis function kernel (RBF) achieves circular shaped areas in the hyperplane for which the configuration parameter γ defines the inverse radius. Additionally, OC-SVM come with the regularization parameter ν defining an upper bound of the accepted outliers and a lower bound for the number of support vectors [25].

We train the OC-SVM with the data of the empty room we later want to run the detector on. Training of a single OC-SVM with data from multiple empty rooms led to low accuracy. However, we used the same hyperparameter for all the different rooms. The optimal hyperperameter configuration for different type of data (RTT and RSSI) and processing chain is shown in Table IV. At the end, the OC-SVM checks whether the feature is within the function (empty room) or if it is an outlier (moving person).

VI. Evaluation

`ToFMD` was prototypically implemented and evaluated by means of experiments in the three indoor scenarios depicted in Figures 9a to 9c.

A. Experimental Setup

As hardware for the WiFi APs we selected ESP32s2 boards (Fig 10) offering a full 802.11n stack operating in 2.4 GHz spectrum and supporting the WiFi FTM protocol. The ESP32s2 boards provide both the RTT value as well as the RSSI value for each exchanged FTM frame which gives

TABLE IV: Hyperparameter selection

data	processing	kernel	γ	ν
RTT	chain A	RBF	0.632412	0.1
RTT	chain B	RBF	0.083417	0.1
RSSI	chain A	RBF	0.498744	0.14
RSSI	chain B	RBF	0.675377	0.11

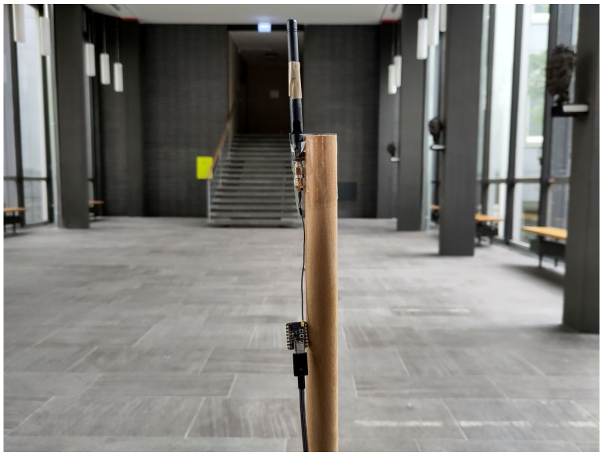


Fig. 10: ESP32s2 setup in the hall of the university building

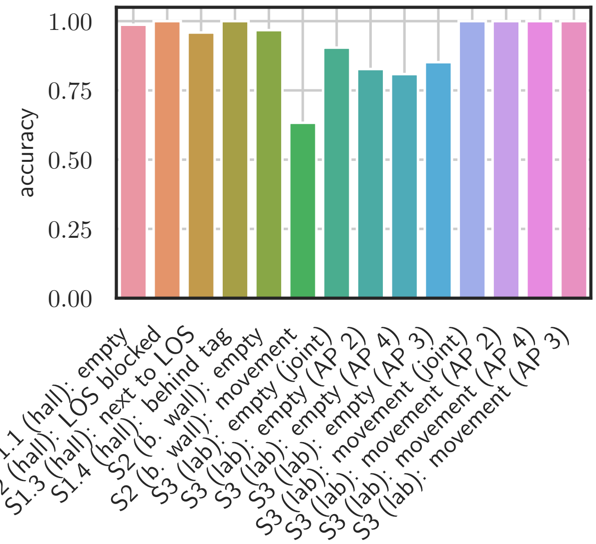


Fig. 11: Accuracy of `ToFMD` using RTT in processing chain A

us 15 values per FTM session send in 2 bursts. The actual mobility detector was implemented in Python. Specifically, we used the implementation *OneClassSVM* of scikit-learn [26] which is based on the work of Schölkopf et al. [27].

We evaluated `ToFMD` in three scenarios:

- *S1*: small hall in our university building (Fig. 9a),
- *S2*: behind a wall in an adjacent room (Fig. 9b),
- *S3*: multi AP setup in our lab (Fig. 9c)

The scenario *S1* is further divided into four sub-scenarios:

- *S1.1*: empty hall
- *S1.2*: walking inside the LOS path
- *S1.3*: walking next to the link
- *S1.4*: walking behind the link

In *S2* we analyzed the performance of `ToFMD` in detection of movements behind a wall. Here, a person was walking close to the other side of the wall. Scenario *S3* allows us to analyze the gain from having multiple anchor nodes from which the RTT/RSSI can be estimated.

B. Performance Metrics

Our main metric for evaluation is the accuracy denoted as:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{7}$$

which takes the true positives (TP), the false positives (FP), true negatives (TN) and false negatives (FN) into account. In our scenarios the room is either empty or somebody is moving inside a given area. Therefore, there are either no TP and FN or no FP and TN. That is why the accuracy is equals the sensitivity in case of a scenario where a person is moving or the accuracy is equals the specificity if an empty room is tested.

C. Results

At first we evaluate the performance of `ToFMD` when using processing chain A. As Fig. 11 shows, the accuracy is above 95% for all sub-scenarios in *S1*. In *S1.2*, which is either a true NLOS or an obstructed LOS (see Fig. 4) scenario, a moving person is detected with 100% accuracy. Scenarios having a LOS condition with additional reflections from the user may have a worse accuracy, as walking in *S1.4* achieves an accuracy of 100% but walking in *S1.3* only achieves an accuracy of 96%. A reason therefore might be the structure of the Fresnel zones which are tighter in the area behind the tag as seen in [9]. The detection of empty rooms has an accuracy of 99% for *S1* and just 90% for *S3*. For *S3* we can see, that fusing the measurements from multiple APs can improve the overall accuracy, as the performance when using only a single AP is at most 85%. An improvement in motion detection when using multiple APs cannot be concluded as already the usage of a single AP achieves 100% accuracy. In *S3* `ToFMD` achieves only an accuracy of 63%. As the wall attenuates the signal, the influence of an reflection on the other side of the wall is small but still measurable.

In a next step we used the RSSI-based approach as a baseline for comparison. While chain A is optimized for processing RTT data we use chain B for a fair comparison as it improves the accuracy when using RSSI data. We do the comparison based on four selected scenarios where `ToFMD` works best (*S1.1* and *S1.2*) but also where it achieves weaker performance e.g., *S1.3* and *S2*.

The results presented in Table V show that if the LOS is blocked or obstructed both approaches achieve an accuracy of 100%. This is probably the easiest case as the signal strength shows strong variations. In all the other cases, where the scenario is a LOS with additional scattering, only the reflection changes which is visible due to constructive or destructive interference. In such cases we see how `ToFMD` outperforms the RSSI approach by 13% in detecting walking next to the link. Also the detection of the empty room seems to raise some issues as the best detection rate using RSSI values is 9% below the accuracy of the RTT approach.

Indeed, the RSSI approach achieves a 20% better accuracy in detecting movements on the other side of the wall using chain B. Chain A using RSSI data was not able to detect any movement. Chain B also improves detecting movement behind the wall for RTT data by nearly 10%. All in all chain A fed with RTT values works better in the other selected scenarios.

VII. Conclusions

In this paper, we proposed `ToFMD` which is a novel technique to detect device-free human movement based on processing ToF (RTT) measurements from COTS WiFi devices. Our solution exploits the influence of multi-path propagation on RTT measurements of COTS WiFi hardware. The approach uses the calculation of the variance as feature extraction and an OC-SVM as detector. The results show an accuracy of more than 80 % and also 100 % in some scenarios like movement in the LOS. `ToFMD` outperforms the baseline of using RSSI values for movement detection. However, we see that detection of movement behind a wall comes with a poor accuracy of only 63 % where using RSSI values and a batch of measurement values works better. Another limit is the room specific training and the careful selection of hyperparameter.

TABLE V: Comparison of `ToFMD` (RTT) with baseline (RSSI).

data type	processing	room	scenario	accuracy
RTT	chain A	hall	S1.1: empty	**98.71** %
RSSI	chain A	hall	S1.1: empty	88.37 %
RTT	chain B	hall	S1.1: empty	87.34 %
RSSI	chain B	hall	S1.1: empty	**89.93** %
RTT	chain A	hall	S1.2: LOS blocked	100.00 %
RSSI	chain A	hall	S1.2: LOS blocked	100.00 %
RTT	chain B	hall	S1.2: LOS blocked	100.00 %
RSSI	chain B	hall	S1.2: LOS blocked	100.00 %
RTT	chain A	hall	S1.3: next to LOS	95.89 %
RSSI	chain A	hall	S1.3: next to LOS	78.54 %
RTT	chain B	hall	S1.3: next to LOS	83.12 %
RSSI	chain B	hall	S1.3: next to LOS	82.65 %
RTT	chain A	b. wall	S2: movement	**63.31** %
RSSI	chain A	b. wall	S2: movement	**0.00** %
RTT	chain B	b. wall	S2: movement	72.40 %
RSSI	chain B	b. wall	S2: movement	**83.12** %

All in all, `ToFMD` can achieve a 10 % improvement compared to an RSSI approach if the setup is placed in the same room. With our approach we enable a better movement detection which is comparable to techniques relying on CSI processing which might not be always available. For future work we plan to extend our work towards a joint detector using RSSI together with ToF data. Additionally, we plan to also include the CSI values in a setup which provides CSI, RSSI and ToF like the ESP32.

Acknowledgements

We would like to thank Alexey Hramov for his contributions in the early phases of the project. This work was supported by the Federal Ministry of Education and Research (BMBF, Germany) within the 6G Research and Innovation Cluster 6G-RIC under Grant 16KISK020K as well as by the German Research Foundation (DFG) within the project ML4WiFi under grant DR 639/28-1.

References

[1] C. Chen, G. Zhou, and Y. Lin, "Cross-Domain WiFi Sensing with Channel State Information: A Survey," *ACM Computing Surveys*, vol. 55, no. 11, pp. 1–37, Feb. 2023.

[2] S. Depatla and Y. Mostofi, "Passive Crowd Speed Estimation in Adjacent Regions With Minimal WiFi Sensing," *IEEE Transactions on Mobile Computing*, vol. 19, no. 10, pp. 2429–2444, Oct. 2020.

[3] H. Huang and S. Lin, "WiDet: Wi-Fi Based Device-Free Passive Person Detection with Deep Convolutional Neural Networks," in *21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2018)*. Montréal, Canada: ACM, Oct. 2018, pp. 53–60.

[4] A. Natarajan, V. Krishnasamy, and M. Singh, "A Machine Learning Approach to Passive Human Motion Detection Using WiFi Measurements From Commodity IoT Devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–10, May 2023.

[5] D. Wu, Y. Zeng, R. Gao, S. Li, Y. Li, R. C. Shah, H. Lu, and D. Zhang, "WiTraj: Robust Indoor Motion Tracking With WiFi Signals," *IEEE Transactions on Mobile Computing*, vol. 22, no. 5, pp. 3062–3078, May 2023.

[6] S. Di Domenico, M. De Sanctis, E. Cianca, and M. Ruggieri, "WiFi-based through-the-wall presence detection of stationary and moving humans analyzing the doppler spectrum," *IEEE Aerospace and Electronic Systems Magazine*, vol. 33, no. 5-6, pp. 14–19, May 2018.

[7] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-Invasive Detection of Moving and Stationary Human With WiFi," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2329–2342, Nov. 2015.

[8] W. Liu, S. Chang, Y. Liu, and H. Zhang, "Wi-PSG: Detecting Rhythmic Movement Disorder Using COTS WiFi," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4681–4696, Mar. 2021.

[9] H. Wang, D. Zhang, J. Ma, Y. Wang, Y. Wang, D. Wu, T. Gu, and B. Xie, "Human Respiration Detection with Commodity Wifi Devices: Do User Location and Body Orientation Matter?" in *ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2016)*. Heidelberg, Germany: ACM, Sep. 2016, pp. 25–36.

[10] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A Real-Time and Contactless Fall Detection System with Commodity WiFi Devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, Feb. 2017.

[11] Y. Zhang, D. Wu, R. Gao, T. Gu, and D. Zhang, "FullBreathe: Full Human Respiration Detection Exploiting Complementarity of CSI Phase and Amplitude of WiFi Signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–19, Sep. 2018.

[12] B. Korany, H. Cai, and Y. Mostofi, "Multiple People Identification Through Walls Using Off-the-Shelf WiFi," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6963–6974, Apr. 2021.

[13] B. Korany and Y. Mostofi, "Counting a Stationary Crowd Using Off-the-Shelf Wifi," in *19th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys 2021)*. Virtual Conference: ACM, Jun. 2021, pp. 202–214.

[14] J. Shang and J. Wu, "A Robust Sign Language Recognition System with Sparsely Labeled Instances Using Wi-Fi Signals," in *14th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS 2017)*. Orlando, FL: IEEE, Oct. 2017.

[15] S. Kianoush, S. Savazzi, V. Rampa, and M. Nicoli, "People Counting by Dense WiFi MIMO Networks: Channel Features and Machine Learning Algorithms," *Sensors*, vol. 19, no. 16, pp. 1–16, Aug. 2019.

[16] Y. Gu, J. Zhan, Y. Ji, J. Li, F. Ren, and S. Gao, "MoSense: An RF-Based Motion Detection System via Off-the-Shelf WiFi Devices," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2326–2341, Dec. 2017.

[17] A. Zubow, K. F. Petto, and F. Dressler, "One-Class Support Vector Machine for WiFi-based Device-free Indoor Presence Detection," in *European Wireless (EW 2023)*. Rome, Italy: IEEE, Oct. 2023.

[18] B. Tan, K. Woodbridge, and K. Chetty, "A real-time high resolution passive WiFi Doppler-radar and its applications," in *IEEE International Radar Conference (RADAR 2014)*, ser. International Radar Conference. Lille, France: IEEE, Oct. 2014, pp. 1–6.

[19] F. Adib, Z. Kabelac, and D. Katabi, "Multi-Person Localization via RF Body Reflections," in *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 2015)*. Oakland, CA: USENIX, May 2015, pp. 279–292.

[20] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Computing Surveys*, vol. 46, no. 2, pp. 1–32, Nov. 2013.

[21] IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE, Std 802.11-2016, Dec. 2016.

[22] S. Lanzisera, D. Zats, and K. S. J. Pister, "Radio Frequency Time-of-Flight Distance Measurement for Low-Cost Wireless Sensor Localization," *IEEE Sensors Journal*, vol. 11, no. 3, pp. 837–845, Mar. 2011.

[23] N. Jathe, M. Lütjen, and M. Freitag, "Indoor Positioning in Car Parks by using Wi-Fi Round-Trip-Time to support Finished Vehicle Logistics on Port Terminals," *Elsevier IFAC-PapersOnLine*, vol. 52, no. 13, pp. 857–862, 2019.

[24] A. Zubow, C. Laskos, and F. Dressler, "Toward the Simulation of WiFi Fine Time Measurements in NS3 Network Simulator," *Elsevier Computer Communications*, vol. 210, pp. 35–44, Oct. 2023.

[25] M. Amer, M. Goldstein, and S. Abdennadher, "Enhancing One-Class Support Vector Machines for Unsupervised Anomaly Detection," in *ACM SIGKDD Workshop on Outlier Detection and Description (ODD 2013)*. Chicago, IL: ACM, Aug. 2013, pp. 8–15.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-Learn: Machine Learning in Python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, Nov. 2011.

[27] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, "Support Vector Method for Novelty Detection," in *12th International Conference on Neural Information Processing Systems (NIPS 1999)*. Denver, CO: MIT Press, Nov. 1999, pp. 582–588.