TagLeak: Non-Intrusive and Battery-Free Liquid Leakage Detection with Backscattered Signals

Conference Paper · October 2018

DOI: 10.1109/ICII.2018.00-16

7 authors, including:

Junchen Guo
Tsinghua University

Meng Jin
Peking University

Chengkun Jiang
Tsinghua University

Yuan He
Tsinghua University

2018 IEEE International Conference on Industrial Internet (ICII)

TagLeak: Non-intrusive and Battery-free Liquid Leakage Detection with Backscattered Signals

Junchen Guo, Ting Wang, Meng Jin, Songzhen Yang, Chengkun Jiang, Long Liu and Yuan He
School of Software and BNRist, Tsinghua University, Beijing, P.R. China
{gjc16, wangt18}@mails.tsinghua.edu.cn, mengj@mail.tsinghua.edu.cn,
{ysz18, jck15, liulong16}@mails.tsinghua.edu.cn, heyuan@mail.tsinghua.edu.cn

***Abstract*—Leakage detection is a crucial issue for factories with numerous pipelines and valves. Conventional methods for leakage detection are mainly rely on manual checking, which results in both high delay and low accuracy. In this paper, we propose TagLeak, a real-time and low-cost system for automatic leakage detection with commercial off-the-shelf (COTS) RFID devices. The key intuition behind TagLeak is that the leaked liquid around tags will change their phase and RSSI (Received Signal Strength Indicator) readings. Multiple challenges need to be addressed before we can turn the idea into a functional system, including: i) it is difficult to detect the slight signal variation that caused by the leaked liquid, based on the coarse-grained RSSI sequence; ii) multipath and interferences can undermine the tags signal, making the variation caused by leaked liquid more difficult to detect.**

We propose solutions to these challenges and evaluate the systems performance in different environments. The experimental results tell that TagLeak achieves a higher than 90.2% true positive rate (TPR) while keeps false positive rate (FPR) below 14.3%. Moreover, as an exploration of the industrial Internet, we have deployed TagLeak in a real-world digital twin system Pavatar for liquid leakage detection in an ultra-high-voltage converter station (UHVCS).

I. Introduction

Liquid leakage is a common problem in todays factories with numerous pipelines and valves. For example, in large-scale rotating machinery like electric generators and synchronous compensators, multiple auxiliary machines for water cooling, lubricant looping, and etc., suffer from potential liquid leakage problems at the unwelded flanges (which is termed as leakage point in this paper), as shown in Fig. 1. Liquid leakage not only affects the normal operation of the machines, but also leads to serious industrial accidents, e.g. circuit shorts due to water leakage, or even fire disasters due to lubricant leakage.

At the heart of addressing this scourge is timely detection. However, todays liquid leakage detection methods still depend on manual checking, e.g., by visual inspections or by touching the surfaces of those machines. Either way can inevitably lead to high detection delay and low accuracy, meanwhile incurs additional labor cost. What's worse, manual dependent methods cannot provide quantized or digitalized monitoring results, which imposes difficulties for further analyses.

There is an urgent need for an automatic leakage detection method. Exploiting water-leakage sensors seems an effective solution, however, due to the limited detection range of the sensors, we have to deploy a large number of sensors to cover the massive and scattered leakage points, which leads to high

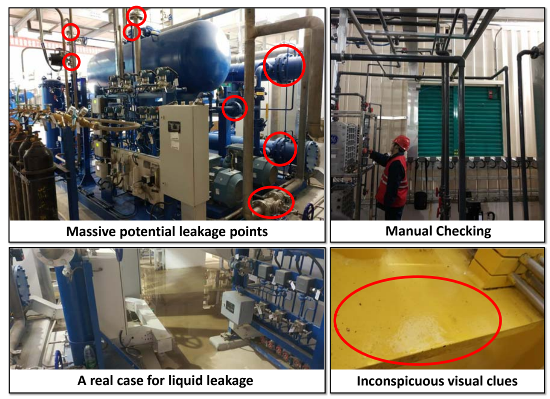


Fig. 1. **Liquid leakage problems in modern factories.**

overhead. Besides, liquid leakage might occur repeatedly and randomly, which requires frequent renewal of the disposable water-leakage sensors. Therefore, water-leakage sensors are seldom adopted in industry due to the unaffordable cost. Camera-based detection may become a low-cost alternative, but it fails in the absence of a line-of-sight(LOS) to the objects or in dark environment.

In this paper, we turn our attentions to a mature technology, RFID (Radio Frequency IDentification). RFID is evolving as a major technology enabler for physical phenomenon sensing, such as vibration inspection [1], eccentricity detection [2], humidity sensing [3], package verification [4], activity recognition [5], touch detection [6] and material identification [7]. The reason for such widespread use of RFID-based sensing are: i) it enables very low lost sensing at high volumes; and ii) it functions in NLOS scenarios and dark environments. This motivates us to think about a challenging question: can RFID-based sensing detect liquid leakage? The answer is yes.

Many works have shown that liquid will change the phase and RSSI of the RF signal. This can be exploited as an indicator for liquid leakage. Though the basic idea sounds straightforward, it is non-trivial to realize this goal due to the following challenges:

- **Challenge 1**: The resolution of RSSI is extremely low. Consider that the signal variation that caused by the leaked liquid is much weaker than the LOS signal at the receiver side, so it is challenging to detect such a slight signal variation based on the coarse-grained RSSI sequence;

- **Challenge 2**: Signal measurement suffers from multiple sources of interferences, such as the vibration of the machine, peoples movements around the tags or reader, the rich multipaths caused by complicated pipelines, etc. These interferences undermine the tags signal, making the variation caused by leaked liquid more difficult to detect.

In this paper, we propose TagLeak, a real-time and low-cost system for automatic leakage detection with commercial off-the-shelf (COTS) RFID devices. TagLeak involves a dual-tag leakage detection device termed as Leakage Detection Tag (LDT), and a Hidden Markov Model (HMM) based leakage detection algorithm. Specifically, LDT leverage the relation between the signals from two tags to filter out the noisy signal introduced by the interferences, and uses the coupling effect between these two tags to enlarge the signal variation caused by liquid leakage. Then a HMM-based detection algorithm is designed to continuously track the variation of the signals from the two tags, which can detect liquid leak with high sensitivity and accuracy.

The contributions of this paper are summarized as follows:

- TagLeak is the first RFID-based system that detect the liquid leakage with high accuracy, using only the coarse-grained backscatter signal. It solves a practical problem for factories with numerous pipelines and valves, which require liquid detection methods that is highly accurate, cost-effective, and robust in different scenarios.
- We design a dual-tag leakage detection device. This makes TagLeak highly sensitive to liquid leakage and at the same time immune to many negative impacts. A HMM-based method is further proposed to detect liquid leakage in real-time.
- We implement a prototype of TagLeak and evaluate it across various scenarios. On average, TagLeak achieves a higher than 90.2% TPR while keeps FPR below 14.3%. As an exploration of the industrial Internet, we have deployed TagLeak in a real-world digital twin system Pavatar [8] [9] for liquid leakage detection in an ultra-high voltage converter station.

The rest of this paper is organized as follows. We introduces some related works of TagLeak in Section II. The preliminary intuition of TagLeak is discussed in Section III. The overview and design details of TagLeak are presented in Section IV. Section V describes the implementation and evaluation of TagLeak. Finally, we conclude TagLeak and discuss future works in Section VI.

II. Related Works

A. Traditional leakage detection methods

Most commodity leakage detection methods exploit water-leakage sensors [10], which detect leakage based on the conductivity of the liquid. These sensors consist of a pair of electrodes and an insulator that separates them. Liquid leakage will cause a short circuit in the two electrodes and then the impedance between them drops sharply, which triggers an alarm. Although this method achieves high detection accuracy, the intrusive deployment and the limited sensing area greatly hampers real-world applications. Compared with the above method, RFID-based system has the advantages of non-intrusive, low-cost, and easy to deploy in various environment.

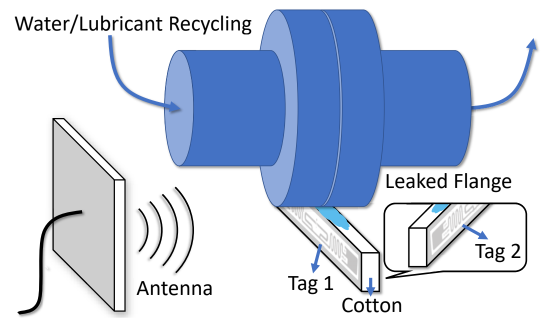


Fig. 2. **Basic experiment setup of TagLeak.**

B. RFID sensing applications

Nowadays, RFID is evolving as a major candidate for cross-modal sensing in industrial scenarios. Specifically, the physical state of the tagged objects as well as their surroundings can be sensed by analyzing the backscatter signals from passive RFID tags. For example, Tagoram [11] and OmniTrack [12] are able to localize and tracking moving objects with high precision. TagBeat [1] and RED [2] are able to detect vibration frequency and eccentricity of a rotating machinery by inspecting tiny signal variations. RIO [6] leverages the fact that the human touch will change the input impedance of a tag's antenna to detect touch actions on critical devices. Similarly, TagLeak discloses special signal patterns for cross-modal sensing, but focuses on another critical task of detecting the liquid leakage.

C. Liquid sensing with RFID

Based on the fact that, as a dielectric, the liquid material can significantly affect the propagation of RF signals by absorbing the electromagnetic fields [13] [14], some works have applied RFID technologies to liquid sensing. For example, TagScan [7] models the signal attenuation when it travels through different kinds of liquid and proposes a distinct feature for material classification. A few works sense ambient humidity with backscattered signals by either exposing the antenna circuit to the vapor [3] or designing specific tags [15]. Different from all the above works, TagLeak leverages the coupling effect for detecting the liquid leakage without manipulating the original COTS RFID systems, and is the first RFID-based leakage-detection system applied into practical industrial scenarios to our knowledge.

III. Leakage Detection With LDT

This section aims to show through both theoretical analysis and experiment that LDT can provide robust signal patterns for the leakage detection.

Fig. 2 shows the sketch of our leakage detection system, which involves a pair of Alien Ultra-High-Frequency (UHF)

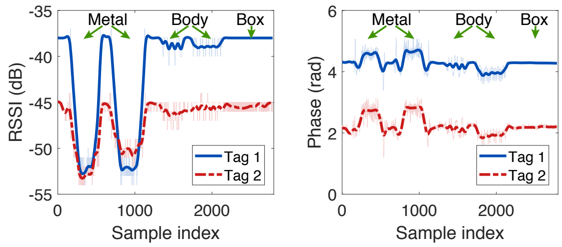


Fig. 3. **Similar signal variation patterns of the two tags under interferences: RSSI (left), Phase (right).**

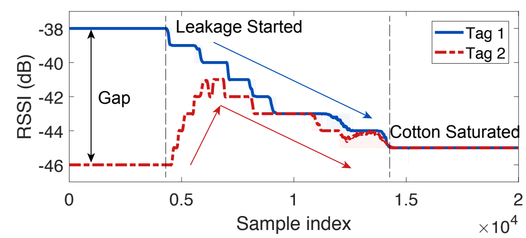


Fig. 4. **Signal variations during the process of liquid leakage.**

passive RFID tags (the reason for exploiting two tags is discussed in Section III-B&C) and a piece of absorbent cotton. The cotton is clamped between the two tags. We term such device as LDT (Leakage Detection Tag). The LDTs are fixed to the flanges. An ImpinJ Speedway R420 RFID reader and a Laird circular polarized antenna are deployed to provide continuous waves and receive the backscattered signals from the tags. When the liquid (e.g. water or lubricant) drops from the unwelded flange, it is absorbed by the cotton which will change the phase and RSSI readings of those two tags. The principle behind is discussed in the following of this section.

A. Absorption Effect of Liquid

Many works have shown that when the backscatter signal penetrates through the medium, different materials of the medium will cause different amounts of phase and RSSI variation of the signal. The amount of variation is mainly determined by the permittivity of the material [13] [14]. In the case of liquid leakage, the leaked liquid, which has higher permittivity than the air, will decrease the strength of the electric field near the tag. This further changes the impedance of the tag's antenna. According to the theoretical transmission model of the RFID signal, the antenna impedance mismatch will not only affect the power charging procedure, but also degrade the signal reflected by the tag, which finally changes the phase and RSSI of the backscatter signal [16]. Moreover, the change in permittivity can also change the wavelength of the backscatter signals, which further incurs different energy loss and phase change along the propagation paths. In summary, we denote the signal degradation induced by the above reasons as the *absorption effect* of liquid.

B. Handling Interference With Dual Tags

Although the phase and RSSI readings provides promising information for leakage detection, we still face a challenge that the signal measurement of the tag suffers from multiple sources of interferences. The main interference comes from the vibration of the rotating machines which the tags are attached to. Moreover, other dynamic interferences e.g. people's movements around the tags or the reader's antenna will also seriously change the backscatter signal.

Fortunately, we find that exploiting the signals from two adjacent RFID tags (as shown in Fig. 2) can help to eliminate the vibration and dynamic interferences. Specifically, since the two tags are attached to the same vibrating target, the micro noise show similar patterns. In the meanwhile, the two closely-deployed tags are expected to face similar dynamic interferences since the distance between them is much smaller than the distance between LDT and dynamic objects.

Fig. 3 shows the signal measurement a LDT attached to a vibrating centrifugal when facing the dynamic interferences induced by metal object, human body and paper box. We find that the irregular vibration noise (light-colored curves) can be filtered out with a low-pass filter and the dynamic interferences dominates the signal variation (dark-colored curves). Besides, the dynamic interferences can be determined by exploring the signal similarity between the two tags.

C. Near-Field Signal Model of LDT

Now, we know that LDT can handle with the interferences, and then we may wonder: What is its signal patterns during the liquid leakage? Can they be distinguished from those during the interferences shown in Fig. 3? In this subsection, we try to answer the above questions by qualitatively modeling the signal of LDT.

Signal Pattern of LDT. We start by introducing the signal patterns of LDT during the liquid leakage (Fig. 4). The patterns can be divided into 3 stages and the characteristics of each stage can be concluded as follows:

1) *Before Leakage*: The signals of the two tags are relatively stable before the leakage. However, there exists a huge gap ($\approx$ 8 dB) between their RSSI readings, although they are very close to each other.
2) *During Leakage*: Tag 2's signal gets stronger at the beginning but gradually decreases as the volume of the liquid increases. On the contrary, Tag 1's signal keeps decreasing during the leakage process.
3) *After Leakage*: When the absorbent cotton becomes saturated, the signals become stable again, and the gap between the RSSI readings almost vanishes in this stage, as shown in Fig. 4.

The above three-stage signal variations of LDT during the liquid leakage show great distinguishability from those during the interferences. The reason behind is explained next.

Qualitative analysis. Fig.4 indicates that the variation of the RSSI gap between Tag 1 and Tag 2 reflects the process of leakage. In the first stage, the initial RSSI gap is induced by the interaction between these two tags. Then, along with the increase of the liquid's volume, such interaction decreases,

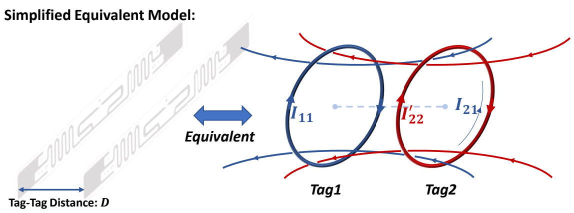
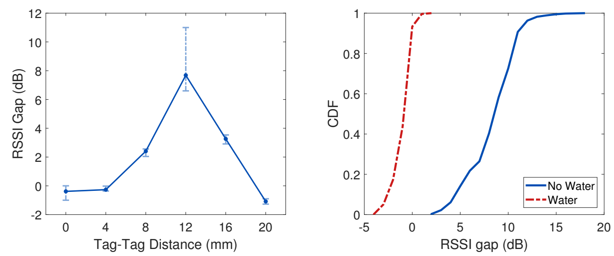


Fig. 5. **Near-field signal model of the LDT.**



(a) Gap V.S. Tag-Tag Distance (b) Gap CDF w/o & w Water

Fig. 6. **Stability of the gap feature.**

and finally vanishes. We qualitatively analyze this procedure in this subsection.

According to the electromagnetic propagation characteristics of small-sized antenna, the radiation field of UHF RFID tag's antenna is classified into near-field induction and far-field radiation. Suppose that the distance from the observation point to the antenna is d, then the boundary of the two fields equals $d^* = \lambda/(2\pi) = c/(2\pi f)$, where λ, f and c are the wavelength, frequency and speed of the electromagnetic wave. Since the tag-tag distance D in a LDT is about 1 to 2 centimeters, which is smaller than the boundary $d^* \approx 5.16$ cm for 925 MHz UHF RFID devices, we should model the near-field coupling effect with the electromagnetic induction model.

We follow the previous works [16] [17] to model a passive tag with a simplified circular loop (Fig. 5), and the current in the circular loop indicates the strength of the backscattered signal. Suppose the original currents in the loops induced by the reader are denoted by I_{11} and I_{22} respectively. When the two adjacent tags interact with each other, the gap of their signal strengths forms mainly due to two effects: *shielding effect* and *coupling effect*.

First, the shielding effect denotes the fact that Tag 1 as a metal object diminish the backscatter communication between Tag 2 and the reader's antenna since Tag 1 is placed in their LOS path (shown in Figure 2). We simply model this signal shielding on Tag 2 by applying an attenuation factor α on its inductive current $I'_{22} = \alpha \cdot I_{22}$, where $0 < \alpha < 1$.

Second, the coupling effect denotes the near-field mutual inductive coupling of the two circular loops. According to the Biot-Savart Law, the inductive current I_{11} in Tag 1 will generate a magnetic field around it, which will further increase the magnetic flux of Tag 2. Then, according to the Lenz's Law, this increasing magnetic flux will induce another inductive current I_{21} on Tag 2, and I_{21} is in the opposite direction of I_{11} to impede the change of the original magnetic flux. The relationship between I_{21} and I_{11} is determined by the mutual impedance. Here we simplify the model by bringing in a scaling factor β, and $I_{21} = -\beta \cdot I_{11}, \beta > 0$, where the negative sign denotes their opposite directions.

Then, we can derive the composite currents in Tag 1 and Tag 2 with the following equations:

$$\begin{aligned} I_1 &= I_{11} + I_{12} = I_{11} - \beta \cdot I'_{22} = I_{11} - \alpha\beta \cdot I_{22} \\ I_2 &= I'_{22} + I_{21} = I'_{22} - \beta \cdot I_{11} = \alpha \cdot I_{22} - \beta \cdot I_{11} \end{aligned} \tag{1}$$

Since the tag-tag distance is far smaller than the reader-tag distance, we assume that $I_{11} \approx I_{22}$. Then, we can qualitatively analyze Tag 1 and Tag 2's signal strengths by comparing their composite currents:

$$\begin{aligned} I_1 - I_2 &\approx (1 - \alpha\beta - \alpha + \beta) \cdot I_{11} \\ &\approx (1 - \alpha)(1 + \beta) \cdot I_{11} \end{aligned} \tag{2}$$

Since $0 < \alpha < 1$, $\beta > 0$, and $(1 - \alpha)(1 + \beta) > 0$, the composite current in Tag 1 is larger than that in Tag 2, then the RSSI gap occurs. From this equation, we can know that the shielding effect $(1 - \alpha)$ leads to the decrease of Tag 2's signal and the coupling effect $(1 + \beta)$ "magnifies" such interaction.

During the leakage process, the interaction between the two adjacent tags as well as the coupling between them and reader's antenna rapidly diminish because of the appearance of the liquid. The liquid with a higher relative permittivity will cause more damage and distortion on electric fields around it, inducing the absorption effect which decreases the signals on both the reader-tag path and the tag-tag path. Therefore, in stage 2 - During Leakage, the RSSI gap vanishes due to the decrease of the interaction between two tags. Besides, Tag 1' signal continuously recedes because the absorption effect dominates the variation, while Tag 2' signal first rises up due to the decrease of the tags' interaction, and then goes down due to the domination of the absorption effect as the increase of the leaked liquid.

D. Opportunity Behind and Its Stability

According to the qualitative analysis above and our preliminary experiments, we find the gap between Tag 1's and Tag 2's RSSI sequences and the tendency of each RSSI sequence are of great use to discriminate the leakage states from the normal states.

To form a distinguishable RSSI gap, we empirically select the tag-tag distance D by practical measurements. We fix the reader-tag distance and change D from $0mm$ to $20mm$. Fig. 6(a) shows the impact of tag-tag distance on the RSSI gap. The near-field interaction reaches maximum when $D = 12mm$. Then we evaluate its stability by conducting experiments with different tags and reader-tag distances in various environment and across various time. Fig. 6(b) shows the cumulative distribution function (CDF) of RSSI gaps without and with the existence of water. We can see that the RSSI gaps show

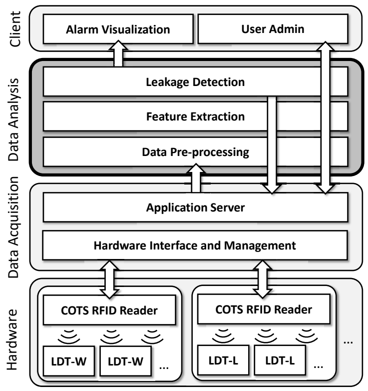


Fig. 7. **TagLeak system overview.**

great difference between the stages before and after the liquid leakage.

IV. TagLeak Design

In this section, we first present a functional overview of TagLeak, followed by details of the technical building blocks.

A. Overview

Fig. 7 presents the system overview of TagLeak which consists of four main modules: hardware, data acquisition, data analysis and client.

- **Hardware**: The hardware of TagLeak consists of multiple COTS RFID readers and multiple LDTs. Each reader covers several LDTs around it. The LDTs are categorized into LDT-W for water leakage detection and LDT-L for lubricant leakage detection.
- **Data Acquisition**: Data acquisition module shields the underlying hardware details, e.g. hardware failure detection and reboot, LDT mapping and data encapsulation, and provides services such as streamed data query, historical alarm query, system configuration and etc.
- **Data Analysis**: Data analysis module is the core of TagLeak. In this module, the RSSI and phase readings are first pre-processed to remove the noises and the interferences. Then, TagLeak extracts features from the processed data and feeds them into a mixed Gaussian-HMM model for timely and accurate leakage detection. Data acquisition module and data analysis module constitute the TagLeak server.
- **Client**: The TagLeak Client visualizes the liquid leakage in a digital twin system. As shown in Fig. 14, the vivid 3-dimension display helps the quick localization of the leakage point. Moreover, it provides the administration service for device and LDT management.

s

Next, we introduced the components of data analysis module in detail.

B. Signal Preprocessing

Raw RSSI and phase data are intrinsically noisy and need processing to improve the accuracy and robustness of further analysis. Before signal processing, we first segment the received signal with a time window W. Length of the window is determined by diffusion rates of the leaked liquid in the absorbent cotton. Consider that different liquid have different diffusion rates in the absorbent cotton which further affects the change rate of the RF signal, W is empirically selected according to the liquid material. After segmentation, the data segments will be pre-processed through the following three steps:

Noise Filtering. The received RSSI and phase signal are susceptible to the variation of the tag which is caused by mechanical vibration. Fortunately, we find that signals variation frequency caused by tags vibration is much higher than that caused by liquid leakage. Therefore, the noise can be filtered out with a Butterworth low-pass filter.

State Estimation. For further data analysis and leakage identification, we first classify the data segment into three states: *stable*, *unstable* and *unreachable*. Segments that belong to different states will be processed with different methods in the subsequence data analysis module. The data is in *stable* state most of the time, but the liquid leakage and the environmental interferences will bring instability to the RF signals, making the signal transfers to the *unstable* state. When the cotton is saturated with the leaked liquid and the undissolved liquid drops adhering to the surfaces of RFID tags, the signal will be absorbed by the liquid, making the signal *unreachable* to the reader.

TagLeak utilizes the RSSI and phase readings for state estimation. Specifically, in the *unreachable* state, the number of samples that the reader received per unit time will decrease significantly. Therefore, the *unreachable* state can be identified using a threshold for the number of samples in a segment. Segments that belong to the *stable* and *unstable* states can be distinguished based on the RSSI variation of the samples in the segment with an empirical threshold. We use a Gaussian mixture model (GMM) to identify the *stable* state, since the average value μ can change over time due to some unpredictable but permanent environmental variations. If a data segment is identified as *stable*, it will be fed into the GMM model for parameter update. Then, μ produced by the GMM model is further leveraged for data normalization.

Interference Determination. As mentioned earlier, the unstability of the data (i.e., the *unstable* state) is caused by liquid leakage or multipath interference. In TagLeak, we would like to filter out the *unstable* segments that caused by multipath interferences. Traditional methods which leverage channel hopping to remove the multipath interference offline is obviously inapplicable for leakage detection [12]. In TagLeak, we leverage the design of *dual*-tag detector (i.e., the LDT) for robust online multipath determination.

As discussed earlier, since the tag-tag distance is small (about 1.4cm in our case), the variation of the signal that caused by the multipath is similar for both the two tags. To verify this assumption, we conduct a set of experiments to observe how different interference affect the signal from the two tags. Fig 3 shows the experimental result. We can see

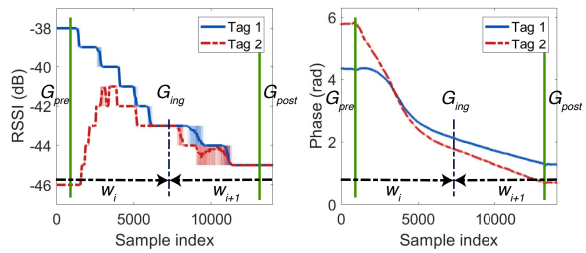


Fig. 8. **Signal variations during the process of liquid leakage.**

that multipath interferences will always result in a consistent trend for Tag 1 and Tag 2. However, if the variation is caused by liquid leakage, the signals from the two tags will exhibit different variation trend, as shown in Fig 4. The reason behind is discussed in Section III-C.

Based on the above result, we propose to identify the interfered segments based on the similarity between the signal from Tag 1 and Tag 2. Specifically, we measure the similarities of the RSSI readings and phase readings from the two tags by calculating their Pearson correlation coefficients, respectively. Features, like kurtosis, spectrum characteristics, variation trend are also considered and fed into a classic CART decision tree. If a segment is deemed as an interference by our classifier, we filter it out first.

C. Feature Extraction

In this subsection, we discuss how to extract features from the processed signals for liquid leakage detection. Before this, we first introduce the variation trend of the signals from the two tags (as shown in Fig. 8) when leakage occurs, which acts as a key feature of liquid leakage. As shown in the figure, the liquid leakage process will go through three stages: pre-leaking stage (termed as G_{pre}), leaking stage (termed as G_{ing}), and post-leaking stage (termed as G_{post}). We have described the signal patterns of each stage in Section III-C.

As the previous description shows, the special variation trend of the signal can be considered as an important indicator about whether liquid leakage occurs. Based on this observation, we propose to consider the similarity between the collected signal (including RSSI and phase) sequence and the pre-constructed signal profile as a feature for leakage detection. Since the leakage speed may be different cross different cases, we propose using Dynamic Time Warping (DTW) to align the collected signal sequence to the signal profile.

However, a problem we meet here is how to accurately capture the G_{ing} stage from the collected signal. We solve this problem leveraging the fact that signal in the G_{ing} stage exhibits much higher variation than that in the G_{pre} and G_{post} stages. Therefore, we detect the starting and ending points of the G_{ing} stage using a moving window (the window here is different from that mentioned in Section IV-C). Specifically, in each window, we calculate the variation of the phase samples in the window (the reason we use the phase samples is that they are more sensitive to the liquid leakage, as shown in Fig. 8). Clearly, the stating and ending points can be identified as the time points where the variation increases and decreases significantly. Then we can capture the G_{ing} stage.

Apart from the variation trend of the signal discussed above, some statistical features like the quantity of readings, the variance sequence of multiple sub-segments and the average distance between the segment, and the stable value , are also extracted for leakage detection. TagLeak extract all these M features from each signal segment. The extracted features form a $M \times K$ feature vector which is fed to the subsequence leakage detection module for further analysis.

D. Leakage Detection Algorithm

As discussed in the previous section, variation trend of the signal acts as a key indicator of whether liquid leakage occurs. That is to say, instead of considering the features of each signal segment independently, we should make a joint consideration of a sequence of signal segments for leakage detection. Therefore, our leakage detection problem is translated to a classification problem that: given a sequence of signal segments and the features for each segment, what is the probability that liquid leakage occurred?

We propose to exploit Markov-based classifier to tackle this problem. In TagLeak, two Gaussian HMM models (denoted by $\mathcal{M}_1$ and $\mathcal{M}_2$) are trained to estimate the probabilities of normal and abnormal events, respectively. Here the normal event means that leakage has not occurred and the abnormal event means that leakage has occurred. $\mathcal{M}_1$ involves two states: G_{pre} and G_{und}, where G_{und} is an undetermined stage, which is possible to transfer to all the other states. $\mathcal{M}_2$ involves all the four states.

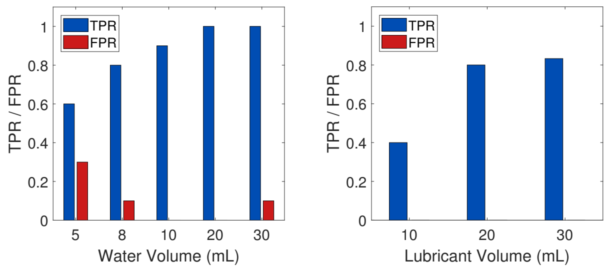
We train $\mathcal{M}_1$ and $\mathcal{M}_2$ exploiting data sets collected from normal event and abnormal event, respectively. Specifically, all the samples in these two datasets are simultaneously fed into $\mathcal{M}_1$ and $\mathcal{M}_2$ to generate a compound representation. For each segment, each HMM model outputs a decoded hidden state sequence of this segment and the logarithmic probability that it belonging to this model (For example, the probability that a segment belongs to $\mathcal{M}_1$ means the probability that this segment is collected when leakage has not occurred). Then, we concatenate the two representations into a vector, and further fed it to a SVM (Support Vector Machine) based classifier, which is trained to determine whether the leakage occurs based on the output vector of the HMM model.

V. Evaluation

A. Implementation

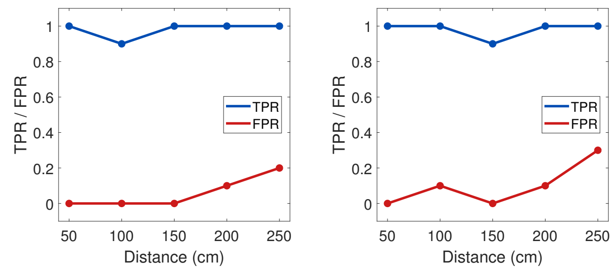
We implement a prototype of TagLeak with COTS ImpinJ R420 RFID reader and our LDT made of a pair of Alien AZ-9640 tags and a piece of absorbent cotton, 10 cm long by 2 cm wide by 1.4 cm by thick.

As we mentioned before, to support the management of multiple readers serving multiple LDTs, we decouple the device management module, the data acquisition module and the data analysis module. The first two modules are implemented



(a) Water Leakage Detection (b) Lubricant Leakage Detection

Fig. 9. **The TPR and FPR over different volumes of different liquid categories.**



(a) Detection with little multipath (b) Detection with rich multipath

Fig. 10. **The TPR and FPR over different Reader-Tag distances in different environments.**

with ImpinJ's application programming interfaces in Java, and are packed into a Websocket socket server to support multiple clients. The encapsulation with Websocket is very suitable for querying streamed data because it eliminates the overhead of constructing multiple HTTP connections while providing more friendly interfaces than transport-layer protocols e.g. TCP. The data analysis module is first implemented in Matlab for a better understanding of the sampled data and a more convenient model exploration with its simulation tools. Then, in the real-world deployment of TagLeak in Pavatar, we optimize the algorithms and pack them into a pure-Java project.

B. Methodology

Since TagLeak focus on the binary classification problem of liquid leakage detection, we mainly use two metrics to evaluate the performance of the proposed approach: true positive rate (TPR) and false positive rate (FPR). TPR represents the percentage that TagLeak correctly detects the liquid leakage. FPR represents the percentage that interferences are mistaken for the liquid leakage by TagLeak.

We mainly discuss the impacts of the following settings on TagLeak's performance:

- **Liquid-related parameter 1 - Volume** (V). The volume of the leaked liquid V reflects the sensitivity of the detection approaches. To evaluate TagLeak, we vary V by 10 mL each time while keeping the duration of the leakage process unchanged.
- **Liquid-related parameter 2 - Category** (C). Since TagLeak is mainly targeted for non-intrusive liquid leakage detection for the water-cooling machines and the lubricant-recycling machines, we evaluate its performances with two categories of liquid: water and lubricant. This experiment evaluates the generalizability of TagLeak.
- **Environment-related parameter 1 - Reader-Tag Distance** (L). We keep using the notion of L to represent the distance between the RFID reader's antenna and LDT(s). To a great extent, L determines the ease of deployment and the coverage of one single antenna. To evaluate this parameter, we move a LDT by 50 cm each time.
- **Environment-related parameter 2 - Multipath and Interference**. These two settings mainly evaluate the practicability of TagLeak system. We choose two different environments with different amounts of multipath. Moreover, we evaluate TagLeak under different kinds of interferences varying in time, location, materials and etc.
- **Algorithm-related parameter 1 - Window Size** (K). K stands for the number of segments in one data window, which is the basic time unit to conduct a prediction in TagLeak. By selecting K, we trade off between the accuracy and the prediction time.
- **Algorithm-related parameter 2 - Classifier**. With the same extracted features, we evaluate different classification methods, e.g. whether they capture the temporal correlation of the segments in one data window.

C. Impact of Liquid Volume

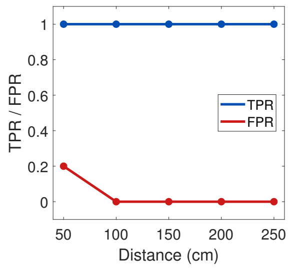
The sensitivity of the detection approaches can be inferred from the relationship between the detection accuracy and the liquid volume V. An approach that achieves higher accuracy over less V is preferred.

We control V by selecting the leaking speed of the medical infusion equipment and remaining a constant duration. The reader-tag distance L is kept 50 cm unchanged. Fig. 9(a) shows the results of water leakage detection over different volume Vs. First, TagLeak can achieve 60% when V is only 5 ml. Then, with the increase of volume, its TPR can reach almost 99%. The TPR over a small quantity of water is acceptable ($>$ 83%) although the pattern of signal variation is less evident. Last, the overall FPR of TagLeak is around 10.0% and the FPR becomes very small (¡ 3.0%) when there is enough liquid absorbed by the cotton. We own this result to the design of LDT, which provides a robust and distinctive pattern while preserves the detection from external interferences.

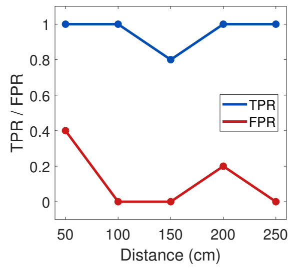
D. Impact of Liquid Category

The generalizability of the detection approaches can be evaluated by applying them to different categories of liquid. Water and lubricant, as the most common as well as critical liquid in modern factories, are taken into consideration.

The same as the settings in the last experiment, we further measures the relationship between the detection accuracy of lubricant and its volume. We can compare the results in Fig. 9 (a) and (b). First of all, the trends are similar: a larger amount

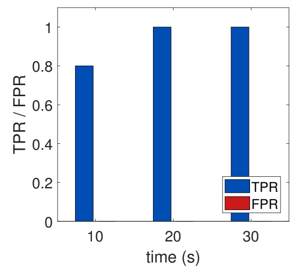


(a) Detection with little interferences

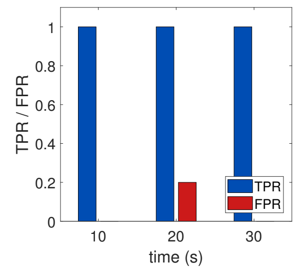


(b) Detection with heavy interferences

Fig. 11. **The TPR and FPR over different Reader-Tag distances under different interferences.**



(a) Interfered by metal



(b) Interfered by human body

Fig. 12. **The TPR and FPR over different interference duration under different material.**

of the leaked liquid lead to more distinctive patterns. However, the TPR of lubricant detection ($> 80\%$) is much smaller than that of water detection, and TagLeak fails to provide reliable predictions when V is relatively small. To illustrate the performance degradation, we conclude the two following reasons: i) lubricant contains less polar molecules and have significantly less absorption of electromagnetic waves than water [14]; ii) the diffusion speed and area of lubricant is much smaller due to its higher viscosity, which further diminishes signal variations.

Thus, to enhance TagLeak with robust lubricant detection, not only the detection algorithm and the quality of the training dataset, but also the design of LDT-L should be improved.

E. Impact of Reader-Tag Distance

One of the advantages of TagLeak is non-intrusive characteristic, because the remote wireless sensing does nearly not block the limited operating space around the auxiliary machines. Thus, the detection range or the reader-tag distance L is a critical metric.

We control L by moving the LDT-W away from the reader's antenna by 50 cm each time, and keep the liquid volume 30 mL. Fig. 10 (a) shows the results of water leakage detection over different tag-reader distance Ls.

The trend of Tag's performance is consistent with our preliminary expectations: As L gets larger, the performance slightly degrades. To achieve the TPR over 95% and the FPR below 5%, we constrain the tag-reader distance no more than 1.5 meter in our practical deployment.

F. Impact of the Amount of Multipath

All of the above basic experiments are conducted in a laboratory with little multipath. However, TagLeak is designed for real-world industrial Internet applications, we evaluates its practicability by deploying the prototype in a rich-multipath environment with many metal tubes.

As shown in Fig. 10 (b), although the performance of TagLeak in the rich-multipath environment is a little poorer, but the average FPR of 10.0% is acceptable in practice. Actually, the degradation is very small, because in a complicated but stable environment, the signal variations are mainly caused by the process of the liquid leakage under LDT's coupling effect.

G. Impact of Interferences

Interferences occurring in the LOS path between the reader's antenna and the LDT can cause significant fluctuations in the RF signals. Fortunately, LDT produces two different patterns of the signal variation when there exists the interference and the leakage, which can be distinguished by our detection algorithm.

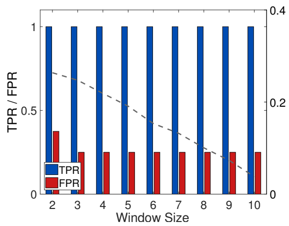
To evaluate the performance of TagLeak, we simulate various levels of interferences by changing the position and duration of the different obstacles. Fig. 11 (a) shows that TagLeak keeps a high TPR ($> 95\%$) along with a low FPR ($< 10\%$), which means TagLeak can effectively tolerate the interferences. With the increase of interferences (Fig. 11 (b)), the TPR of TagLeak remains almost unchanged, but the FPR goes up slightly. Moreover, when the tag-reader distance is small, the FPR is relatively higher because the similarity of the signal variations of two tags slightly goes down due to the increase of the relative magnitude of tag-tag distance D and tag-reader distance L.

Then, we evaluate the performance of our interference determination algorithm. A metal plate and one hand of an author are used as the obstacles blocking the LOS path of the LDT. Fig. 12 (a) and (b) shows the results respectively. The TPR means the accuracy that the interferences are recognized correctly and the FPR stands for the probability that the sequences during the liquid leakage are falsely deemed as interferences. The results show that our interference determination algorithm in TagLeak achieves overall TPR around 96.7% and overall FPR around 3.3%. It can effectively recognizes the interferences regardless of the interference time and materials (moving metal and human body).

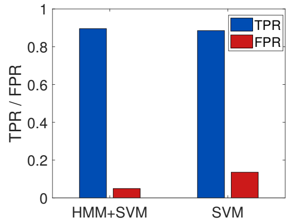
H. Impact of Model Parameters

The design of the detection algorithm can also significantly affect the performance. Here, we mainly focus on two key components: the detection window size K and the backend classifier.

Window size K determines the length of time we take to collect data, a smaller K stands for a higher level of sensitivity for the leakage detection. Fig. 13 (a) shows that K has relatively small effect on the accuracy of TagLeak. However, with the increase of K, the processing delay is



(a) Window size and performance



(b) Classifier and performance

Fig. 13. **The TPR and FPR over different settings of different model parameters.**

reduced, because less data windows are generated. To balance the sensitivity and the system overhead, we set K equals to 5 in the other experiments.

Fig. 13 (b) shows the comparison of two backend classifier. The HMM+SVM approach is TagLeak's current choice and the pure SVM are fed in with whole features without any representation conversion. We evaluate the two models with the whole test dataset consisting of multiple samples collected in various scenarios. From the figure, we know that both of the models achieve relatively high TPR, 90.1% and 88.9% respectively. But TagLeak's approach achieves lower FPR at about 4.9%. We own this improvement to the import of HMM. Because it considers more about the temporal correlation between adjacent segments, which can tell from the differences between a leakage sequence and an interference sequence that the fluctuations in the latter one will disappear and the signal will come back to G_{pre}.

VI. Conclusion and Future Works

In this paper, we advance the state of the arts by applying the RFID battery-free sensing technology to a widespread problem, liquid leakage detection in smart factories and plants. The proposed system TagLeak first extracts robust features of the signal model based on the coupling effect of two adjacent tags, and then detects the liquid leakage from pre-processed data stream with a HMM-based classification algorithm. Our solution is not only non-intrusive and low-cost in deployment, but also real-time and robust in practical use. The experiment results tell that TagLeak achieves a higher than 90.2% TPR while keeps FPR below 14.3%. As an exploration of the industrial Internet, we deploy a prototype of TagLeak in a digital twin system Pavatar in Hunan UHVCS [8] [9] to monitor 38 potential leakage points. Fig. 14 shows the real-world deployment and corresponding 3D visualization of leakage warnings. The prototype has been running for 3 months before the submission of this paper.

Although TagLeak has made a step towards industrial IoT, there still have many corresponding future works. For example, conducting fined-grained sensing with leakage volume and rate estimation, improving the detection accuracy of low-permittivity materials, e.g. lubricant, expanding the coverage and density of LDT deployment, exploring the parallel sensing capability and etc., are all critical challenges to be solved.

Fig. 14. **Real-world deployment and visualization of TagLeak in Pavatar.**

VII. Acknowledgment

This work is sponsored by National Key R&D Program of China No. 2017YFB1003000, National Natural Science Foundation of China No. 61772306, and State Grid of China Research Fund.

References

[1] Lei Yang, Yao Li, Qiongzheng Lin, et al. Making sense of mechanical vibration period with sub-millisecond accuracy using backscatter signals. In *Proceedings of ACM MobiCom*, 2016.

[2] Yilun Zheng, Yuan He, Meng Jin, et al. Red: Rfid-based eccentricity detection for high-speed rotating machinery. In *Proceedings of IEEE INFOCOM*, 2018.

[3] S Manzari, C Occhiuzzi, S Nawale, et al. Polymer-doped uhf rfid tag for wireless-sensing of humidity. In *Proceedings of IEEE RFID*, 2012.

[4] Ge Wang, Chen Qian, Jinsong Han, et al. Verifiable smart packaging with passive rfid. In *Proceedings of ACM Ubicomp*, 2016.

[5] Longfei Shangguan, Zimu Zhou, Xiaolong Zheng, et al. Shopminer: Mining customer shopping behavior in physical clothing stores with cots rfid devices. In *Proceedings of ACM Sensys*, 2015.

[6] Swadhin Pradhan, Eugene Chai, Karthikeyan Sundaresan, et al. Rio: A pervasive rfid-based touch gesture interface. In *Proceedings of ACM MobiCom*, 2017.

[7] Ju Wang, Jie Xiong, Xiaojiang Chen, et al. Tagscan: Simultaneous target imaging and material identification with commodity rfid devices. In *Proceedings of ACM MobiCom*, 2017.

[8] Pavatar project. http://tns.thss.tsinghua.edu.cn/sun/pavatar.html, 2018.

[9] Yuan He, Junchen Guo, and Xiaolong Zheng. From surveillance to digital twin: Challenges and recent advances of signal processing for industrial iot. *IEEE Signal Processing Magazine*, 2018.

[10] Honeywell Inc. Leakfilm product. https://www.honeywellanalytics.com/en/products/LeakFilm, 2018.

[11] Lei Yang, Yekui Chen, Xiang-Yang Li, et al. Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices. In *Proceedings of ACM Mobicom*, 2014.

[12] Chengkun Jiang, Yuan He, Xiaolong Zheng, et al. Orientation-aware rfid tracking with centimeter-level accuracy. In *Proceedings of ACM/IEEE IPSN*, 2018.

[13] Daniel M Dobkin and Steven M Weigand. Environmental effects on rfid tag antennas. In *IEEE MTT-S International Microwave Symposium Digest*, 2005.

[14] Joshua D Griffin, Gregory D Durgin, Andreas Haldi, et al. Rf tag antenna performance on various materials using radio link budgets. *IEEE Antennas and Wireless Propagation Letters*, 2006.

[15] Juha Virtanen, Leena Ukkonen, Toni Bjorninen, et al. Inkjet-printed humidity sensor for passive uhf rfid systems. *IEEE Transactions on Instrumentation and Measurement*, 2011.

[16] Klaus Finkenzeller. *RFID handbook: fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication*. John Wiley & Sons, 2010.

[17] Ge Wang, Chen Qian, Haofan Cai, et al. Replay-resilient physical-layer authentication for battery-free iot devices. In *Proceedings of ACM HotWireless*, 2017.