TwinLeak: RFID-based Liquid Leakage Detection in Industrial Environments

Junchen Guo*, Ting Wang*, Yuan He*, Meng Jin*, Chengkun Jiang* and Yunhao Liu*†
*School of Software and BNRist, Tsinghua University, †Computer Science and Engineering, Michigan State University
{gjc16, wangt18}@mails.tsinghua.edu.cn, {heyuan, mengj}@mail.tsinghua.edu.cn,
jck15@mails.tsinghua.edu.cn, yunhao@cse.msu.edu

***Abstract*—Liquid leakage detection is a crucial issue in modern industry, which concerns industrial safety. Traditional solutions, which generally rely on specialized sensors, suffer from intrusive deployment, high cost, and high power consumption. Such problems prohibit applying those solutions for large-scale and continuously industrial monitoring. In this work, we present a RFID-based solution, TwinLeak, to detect liquid leakage using COTS RFID devices. Detecting the leakage *accurately* with coarse-grained RSSI and phase readings of tags has been a daunting task, which is especially challenging when *low detection delay* is required. Our system achieves these goals based on the fact that the *inductive coupling* between two adjacent tags is highly sensitive to the liquid leaked between them. Therefore, instead of judging according to the signals of each individual tag, TwinLeak utilizes the relationship between the signals of two tags as an effective feature for leakage detection. Specifically, TwinLeak extracts discriminative signal features from short segments of signals and instantly identifies leakage using a light-weight classifier. A model-guided method for leakage progress tracking is further devised to simultaneously estimate the leakage volume and rate. We implement TwinLeak, evaluate its performance across various scenarios, and deploy it in a real-world industrial IoT system. In average, TwinLeak achieves a TPR higher than 97.2%, a FPR lower than 0.5%, and a relative property estimation error around 10%, while triggering early alarms after only about $4.6mL$ liquid leaks.**

I. Introduction

In modern industry, auxiliary machines used for water cooling, lubricant looping, and liquid purification are indispensable. Fig. 1 shows a real industrial scenario, where a number of flanges connect pumps and tanks. The flanges are often not welded for convenient maintenance and replacement, potentially leading to the liquid leakage problem. Liquid leakage is indeed a frequent accident that threatens the safety (e.g. fire disasters caused by leakage of combustibles) and incurs countless economic loss (e.g. machine damage caused by leakage of cooling water or lubricant). Liquid leakage detection is therefore an extremely significant task in the operations of modern industry.

Conventional leakage detection still relies on the labor-intensive, high-delay, and inaccurate manual checking. Many efforts have been taken to achieve automatic leakage detection, but the existing solutions remain inapplicable in complex and large-scale industrial scenarios. For instances, vision-based approaches fail to work in dark environments or the none-line-of-sight (NLOS) scenarios, which are common in industrial environments. Detection with specialized sensors seems an effective alternative, but most sensors suffer high cost of deployment and maintenance, high power consumption, and difficulty of deployment into many machines.

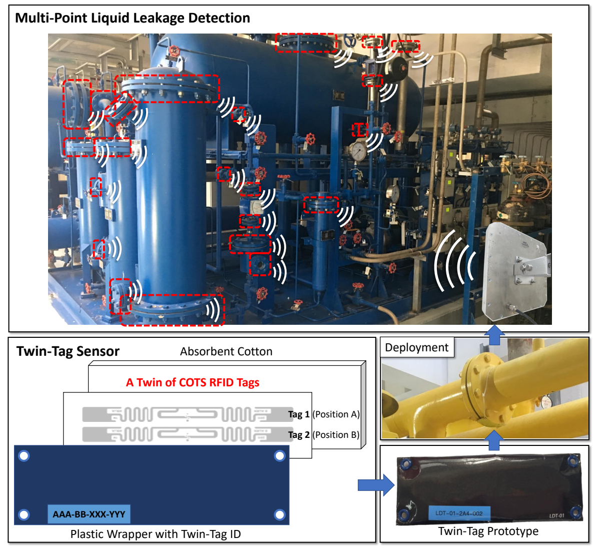


Fig. 1. **Liquid Leakage Detection with Twin-Tag.**

Motivated by the recent research advances in RFID-based sensing ([1] [2] [3] [4] [5] [6]), in this work we study the problem of RFID-based liquid leakage detection in industrial environments. Recent related works (e.g., exploiting RF signal for liquid identification [5] [7] and humidity sensing [8] [9]) have shown that the existence of liquid may change the characteristics of the RF signals, which might be exploited as indicators for liquid leakage. But the liquid leakage detection problem in industrial environments differs from the existing works, due to the following critical requirements and challenges:

- **Timeliness requirement.** To minimum the damage caused by liquid leakage, we must detect the the leakage event in a timely manner, i.e. once a few drops of liquid leak out. The corresponding signal variation in this case, however, is too slight to be detected by COTS RFID devices. Therefore, existing methods which identify liquid based on the obvious signal variations caused by more than 100 mL of liquid [5] [7] is not applicable in the leakage detection scenario.
- **High accuracy requirement.** Our goal is to detect the leakage events that occur exactly at the potential leakage points (e.g., the flanges). However, many other confound-

ing events (such as the change in the indoor humidity, or the dynamics in the surrounding environment) affect the signal propagation between the tags and the reader. Those events also change the characteristics of the signals, leading to false alarms. Excessive false alarms are an equally serious problem, because they will incur extra labor cost to deal with the alarms, cause unnecessary downtime of the machinery, or even lead to inappropriate operations.
- **High precision requirement.** Besides accurate detection, quantified properties of the leakage process (such as the leakage rate and volume) also act as important decision-making basis when leakage occurs. It is difficult to obtain such fine-grained information from the coarse-grained RSSI and phase readings.

In this paper, we present TwinLeak, a ready-for-use solution to accurately and timely detect liquid leakage and quantify the leakage rate and volume using COTS RFID. Our idea is based on the *inductive coupling effect* between two adjacent tags. We find through extensive experiments that even slight amount of liquid may change the dielectric between two adjacent tags, which leads to significant variation in the strength of inductive coupling. Hence, observing the relationship between the signals of two tags offers rich information to detect and quantify the liquid leakage process. We address non-trivial challenges in implementing the above idea and make TwinLeak a practical solution. The contributions of this work are summarized as follows:

- We disclose and model the relationship between liquid leakage and the inductive coupling effect between tags. Instead of judging according to the signals of each individual tag, we identify the relationship between the signals of two tags as an effective feature for leakage detection.
- We propose the TwinLeak approach, which addresses several unique technical challenges in leakage detection. (i) We extract discriminative signal features and design a decision tree based classification method for instant leakage detection; (ii) We provide a change point detection algorithm to accurately localize the starting points of the leakage; (iii) We design a model-guided method for leakage progress tracking, which simultaneously estimates the leakage properties e.g. volume and rate.
- We implement TwinLeak on COTS RFID, deploy it in a real-world industrial monitoring system (as shown in Fig. 1), and evaluate its performance across various scenarios. TwinLeak achieves a TPR higher than 97.2%, a FPR lower than 0.5% and a relative property estimation error about 10%, with the early alarms after only about $4.6mL$ liquid leaks.

The rest of this paper is organized as follows. We discuss the related works of TwinLeak in Sect. II. Sect. III introduces the Twin-Tag sensor design and its intuition. The overview and design details of TwinLeak are described in Sect. IV and Sect. V&VI respectively. Sect. VII presents the implementation and evaluation of the system. Finally, Sect. VIII summarizes TwinLeak and discusses potential research extensions.

II. Related Works

In this section, we give a brief review of the related works of TwinLeak. They are mainly categorized into three groups.

A. Liquid leakage detection

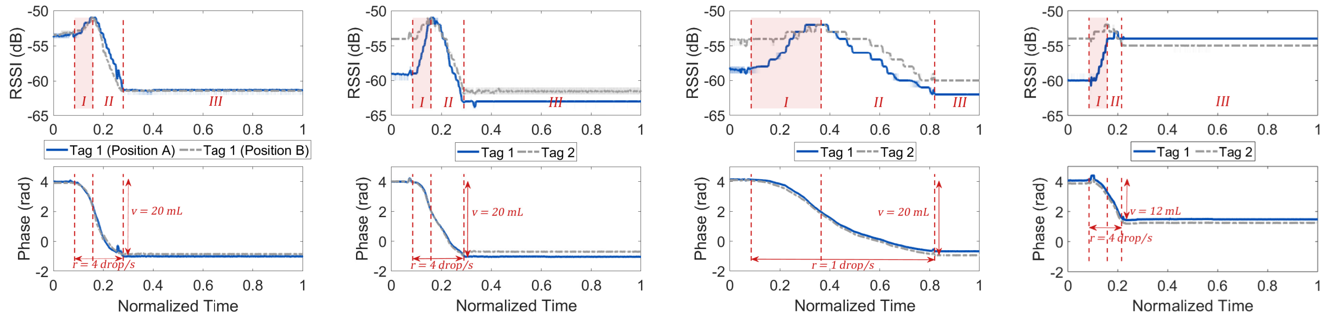
Many efforts have been made to detect liquid leakage in both research and application fields. Various sensors based on acoustic/vibration [10], pressure [11] and electrode [12] are proposed to achieve the high accuracy and sensitivity. However, most of them suffer from the issue of intrusive-deployment and the high maintenance cost. Although acoustic sensors provide a promising coverage, they are not suitable for extremely noisy environment in most of the industrial environments. TwinLeak outperforms these methods for its non-intrusive, battery-free, low-cost and easy-to-deploy characteristics as well as a relatively large wireless sensing range.

B. RF-based liquid detection and identification

Wireless signals are proved susceptible to the materials of the medium that along the propagation path [5] [7] and near Tx/Rx antennas [8] [9] [13], which provides an opportunity for sensing the characters of the liquid. By modeling the relationship between the signal attenuation and the permittivity of the propagation medium, TagScan [5] and LiquID [7] achieve liquid identification / classification with CTOS RFID devices and UWB devices respectively. However, a large amount of liquid (usually $> 100mL$) is required to block the propagation path, which greatly restricts the practical deployment. Recent humidity sensing approaches have demonstrated that the near-field coupling of the liquid and the antenna offers much better sensitivity. However, existing approaches need either high-overhead operations, e.g. the *power-on* measurement under different frequencies [8], or specially-designed tags to magnify corresponding signals [9] [13]. Compared with the above mentioned methods, TwinLeak is able to detect just small amount of leaked liquid at specific points by just analyzing the RSSI and phase readings of COTS RFID tags.

C. Exploration of coupling effect

Although the coupling effect between tags has been considered as detrimental in many existing RF-based sensing methods, many works exploit such phenomenon for more accurate sensing. For instance, Twins [14] tracks an object by observing the variation of the coupling effect between two adjacent tags attached on its surface. HuFu [15] proposes that the coupling effect indeed creates a distinct fingerprint for the signal from two adjacent tags, which offers an opportunity to resist the replay attacks. RIO [4] exploits the coupling effect to sense and track human's touch on a multi-tag array. Our work leverages the coupling effect for liquid leakage detection. We observe that the coupling effect can not only enlarge the signal variation that caused by the liquid leakage, but also filter out the variation that caused by other multipath interferences. This



(a) Quick Leakage (Single) (b) Quick Leakage (Twin) (c) Slow Leakage (Twin) (d) Unsaturated Leakage (Twin)

Fig. 2. **Signal Patterns of Twin-Tag.**

provides an opportunity to achieve accurate and timely leakage detection.

III. Liquid Leakage Primitive

In this section, we start with a introduction about the design of our Twin-Tag sensor, followed by the empirical observation and theoretical explanation about Twin-Tag's distinct signal for the leakage detection.

A. Twin-Tag sensor

Fig. 1 shows the structure of the Twin-Tag sensor, which consists of a pair of parallel and proximate COTS RFID tags, a piece of absorbent cotton and a plastic wrapper. The twin tags are attached to the cotton, which acts as a container to collect the leaked liquid. Moreover, it separates the twin tags from the monitored target (e.g., the flanges, which are generally metal) to avoid the electromagnetic shielding. A piece of plastic film wraps the twin tag and the cotton, which excludes the disturbances from external environmental factors, e.g. humidity. Twin-Tag can be easily pasted or tied to the leakage points (Fig. 1). When the liquid starts to leak out, the cotton absorbs the liquid, which influences the RSSI and phase of the sensor's signal.

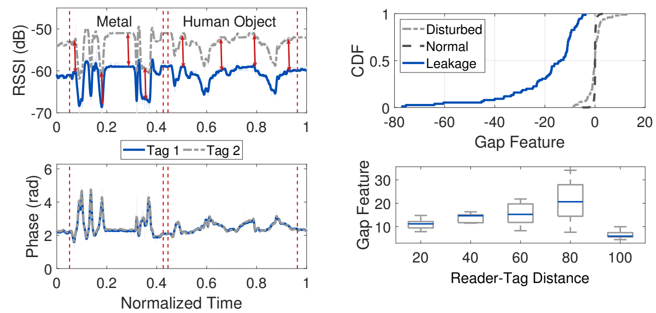
The core intuition of Twin-Tag's design is the integration of two parallel and proximate tags. The reasons we use a pair of tags instead of a single tag are:

1) Signal variation of a single tag is too slight to be stably detected in the early stage. The mutual inductive coupling of the twin tags can amplify such variation.
2) Inductive coupling effect between two adjacent tag can also filter out the variation that cased by other confounding events.

B. How does leakage change the signal of Twin-Tag?

Figure 2 demonstrates the backscatter signal of Twin-Tag that collected during leakage processes in various scenarios.

(1) Signal of a single tag. We first explore how the leaked liquid affects the signal of a single tag. Specifically, Fig. 2(a) shows the RSSI and phase measurements obtained when a single tag is located at different positions (termed as Position A and B shown in Fig. 1). We observe that (i) the signals collected at different locations are almost identical, indicating that tag's signal is almost irrelevant to its location in the sensor. (ii) The RSSI and phase measurements remain stable before the leakage. (iii) After the leakage starts, the RSSI rises slightly at the beginning (termed as *Stage I*), then drops with the increased amount of the liquid (termed as *Stage II*), and finally becomes stable again (termed as *Stage III*) when the cotton is saturated ($\sim 20mL$). Meanwhile, the phase value decreases monotonously.



(a) Confounding Interferences (b) Validation of Gap Feature

Fig. 3. **Interference signals and the robust feature of RSSI gap's trend.**

(2) Signal of Twin-Tag. Now we show that Twin-Tag offers a distinct evidence for the early-stage leakage detection due to the the interplay between the two tags (Fig. 2(b)). Specifically, we find that (i) before the leakage, there is a gap between the RSSI measurements of the two tags, although they suffer from similar reader-tag path loss. (ii) When the leakage starts, the RSSI of Tag 1 rises sharply, and hence the RSSI gap gets narrowed rapidly in *Stage I*.

The above observation brings dual benefits to the early-stage leakage detection: (i) Compared to the single-tag design, the Twin-Tag offers the distinguishable pattern in *Stage I*. (ii) And such pattern can remarkably reduce the false alarms caused by other confounding events, since the closely-deployed twin tags experience similar interferences from other confounding events.

Fig. 3(a) shows signal of the Twin-Tag collected when obstacles (metal objects and human bodies) move around the reader-tag LOS path. We can see that the RSSI gap stays almost unchanged even during rapid fluctuations while the phase readings of the two tags are almost identical. Therefore,

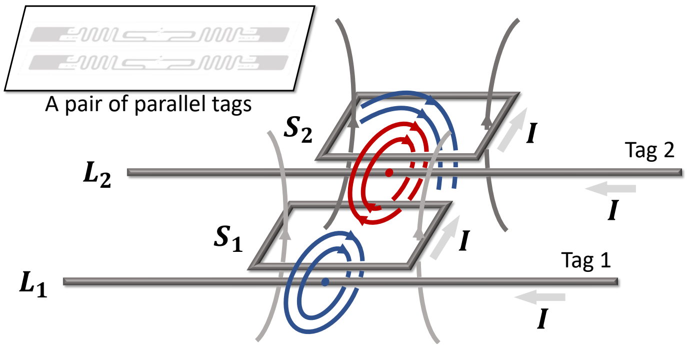


Fig. 4. **Asymmetric Mutual Inductive Coupling Effect.**

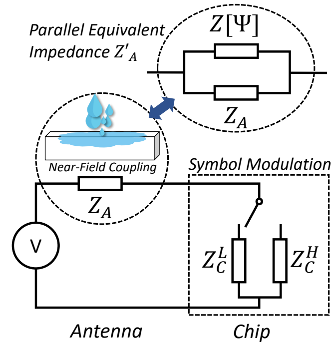


Fig. 5. **RFID tag Model.**

we can define a robust gap feature to recognize *Stage I* during leakage:

$$f = \frac{d(r_2 - r_1)}{dt} \cdot \frac{dr_1}{dt} \quad (1)$$

where r_1, r_2 stands for the RSSI readings of Tag 1 and Tag 2 respectively. Fig. 3(b-Top) shows the CDF of the gap feature in various conditions, where we can conclude that the early stage of the liquid leakage is distinguishable from normal state or other confounding events. By fixing the leakage rate, we further verify the robustness of the gap feature to distance, as shown in Fig. 3(b-Bottom).

(3) Impact of the leakage rate and volume. We also observe the impact of the leakage rate and volume to the signal pattern of the Twin-Tag. As shown in Fig. 2(c)(d), we find that: (i) The leakage rate won't change the shape of the signal pattern, but will only induce temporal stretching and shrinking. (ii) The leakage volume only cause the spatial truncation without changing the shape either. Thus, the ability of early-stage detection of Twin-Tag preserves, and in the meanwhile, an opportunity is offered to estimate extra properties of the leakage process.

C. Theoretical analysis of the signal pattern

Next, we explain the signal pattern of the Twin-Tag.

(1) The initial RSSI gap. We notice that there exists an initial RSSI gap between Tag 1 and Tag 2 although they suffer from similar reader-tag path loss. Since such phenomenon does not preserve in the single-tag measurement, we attribute it to the asymmetric interaction called *mutual inductive coupling* between the two adjacent tags, which results from the near-field electromagnetic induction [14] [15]. Next, we exploit the structure-aware and asymmetric model introduced by Twins [14] to explain the initial RSSI gap of the Twin-Tag.

As shown in Fig. 4, the dipole antenna is deemed to be equivalent to an electric dipole (denoted as a line L) and a magnetic dipole (denoted as a rectangle S) in this model [14]. The equal currents induced by the reader on L_i and $S_i, i \in \{1, 2\}$, will generate a magnetic field around themselves, and further mutually change the magnetic flux in S_1 and S_2:

$$\begin{aligned}\Phi^{\Sigma}_{S_1} &= +\Phi_{S_1} - \Phi_{L_1 \to S_1} - \Phi_{S_2 \to S_1} + \Phi_{L_2 \to S_1} \\ \Phi^{\Sigma}_{S_2} &= +\Phi_{S_2} - \Phi_{L_2 \to S_2} - \Phi_{S_1 \to S_2} - \Phi_{L_1 \to S_2}\end{aligned} \quad (2)$$

where $\pm$ stands for the direction of the flux, Φ represents its absolute value. Due to the symmetry of S_1 and S_2 to the reader, we assume that the original fluxes (Φ_{S_1} and Φ_{S_2}), and their mutual fluxes ($\Phi_{S_2 \to S_1}$ and $\Phi_{S_1 \to S_2}$) are identical. Besides, the fluxes induced by adjacent L_i ($\Phi_{L_1 \to S_1}$ and $\Phi_{L_2 \to S_2}$) are also identical. However, the fluxes induced by farther L_i are totally different. According the Ampere's right-hand rule and the Biot-Savart Law, they are in the opposite directions and $\Phi_{L_2 \to S_1} > \Phi_{L_1 \to S_2}$. Therefore, according to the Faraday's law of electromagnetic induction, the inductive electromotive force of Tag 1 is smaller than Tag 2, which explains the initial RSSI gap:

$$\Delta \mathcal{E} = -\frac{d}{dt}(\Phi^{\Sigma}_{S_2} - \Phi^{\Sigma}_{S_1}) = \frac{d}{dt}(\Phi_{L_1 \to S_2} + \Phi_{L_2 \to S_1}) > 0 \quad (3)$$

(2) Signal variation during leakage. Since the leaked liquid does not block the reader-tag LOS path, the backscattered signal variates due to the near-field coupling between the liquid and the tag's antenna. We model the liquid with an equivalent impedance $Z(\Psi)$, where Ψ relates to its volume. Then, the tag's input impedance Z_A becomes $Z'_A(\Psi) = \frac{Z_A \cdot Z(\Psi)}{Z_A + Z(\Psi)}$. Denoting the tag's chip impedance as Z_C, now we can note that the power-reflection coefficient K which represents the ratio of backscattered power becomes:

$$K(\Psi) = |1 - \rho(\Psi)|^2 = \left|1 - \frac{Z_C - Z^*_A(\Psi)}{Z_C + Z_A(\Psi)}\right|^2 \quad (4)$$

Along with the increasing of Ψ, the impedance matching relationship $Z_C = Z^*_A(\Psi)$ no long preserves, and the received signal at the reader side becomes weaker. We attribute the slight RSSI rise of a single tag in *Stage I* (Fig. 2(a)) to the initial imperfect impedance matching [16].

However, such pattern is too slight to be robustly detected, which motivates the design of the Twin-Tag: By considering the mutual coupling during the leakage, we "magnify" the signal variation in *Stage I* for early-stage detection. When the liquid leakage occurs, the mutual inductive coupling gets weak, and thus $\Delta\mathcal{E}$ gradually drops to zero. Therefore in *Stage I*, the RSSI value of Tag 1 rises rapidly closer to Tag 2, which contributes a distinct and detectable signal pattern.

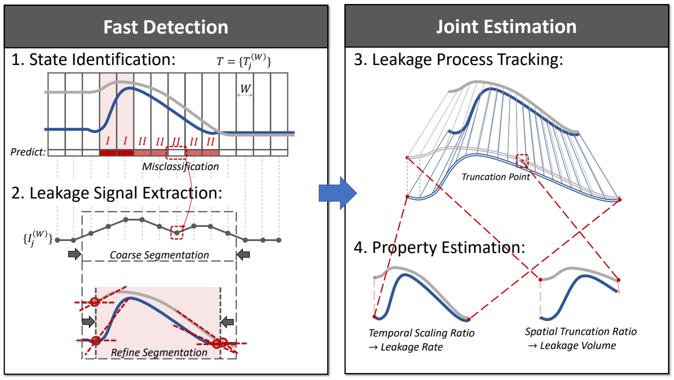


Fig. 6. **The outline of data processing in TwinLeak.**

Feature
Variance
Deviation From "Stable"
Linear-Fit Slot
Linear-Fit RMSE
Diff. Variance
Diff. Average
Gap Feature

Fig. 7. **Feature Table.**

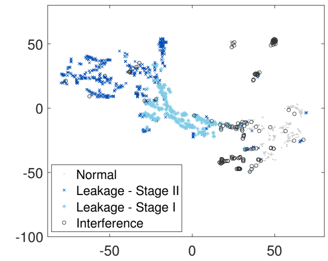


Fig. 8. **Clustering Visualization.**

IV. TwinLeak Overview

In this work, we leverage the above characteristics of the signal and propose TwinLeak, a low-cost system that is able to instantly and accurately detect liquid leakage using only the COTS RFID devices. Fig. 6 shows its processing outline, in which we design two components: *fast detection* and *joint estimation*.

(1) Fast Detection. To achieve real-time detection, TwinLeak uses a short time window to collect the RSSI and phase readings, and identify whether leakage occurs based on the extract signal features in the current time window. Here we empirically set the size of the time window as $10s$. Once successive windows are determined as the *leakage state* (which means a high probability for leakage), TwinLeak gives an alarm (Sect. V-A) and extracts corresponding signal segments for further analyses (Sect. V-B).

(2) Joint Estimation. In addition to the detection of leakage, some other fine-grained information, such as the probability of leakage, the leakage rate, and the leakage volume are also the concerns of the industrial fields. To address this problem, TwinLeak tracks the progress of leakage by matching the extracted signal segments with priori templates. Specifically, we introduce a model-guided leakage tracking algorithm that performs fast signal matching, and at the same time tolerates local jitters while preserving signal variation patterns induced by the leakage (Sect. VI-A). The matching method provides two information: i) the distance between the extracted signal segments and the priori templates, which can be used for leakage probability estimation; ii) the progress of the leakage process, which can be used for leakage rate and volume estimation (Sect. VI-B).

V. Fast Detection

The gap feature introduced in Sect. III-B gives us an opportunity to trigger an alarm for the leakage event in the early stage. We call this component *fast detection*, which is done in two steps: *state identification* and *leakage signal extraction*.

A. State identification

The target of state identification is to mark the signal in each window with a certain state: *normal state*, *disturbed state*, or *leakage state*, where the disturbed state means the signal variation that caused by confounding events. The intuition behind our state identification method is the signal pattern of *leakage state*, which is the superimposition of the mutual inductive coupling between tags and the impedance changing caused by the leaked liquid, is distinct from the other states. Therefore, we first extract the features that can distinguish different states of each time window, and then identify the state based on those features.

Fig. 7 shows a set of selected features. As shown in the figure, except from the gap feature, other statistic features and trend features are considered for both the signal of each individual tag and the differential signal of the two tags for accurate classification. Fig. 8 further visualizes the clustering of the extracted features with t-SNE (t-distributed Stochastic Neighbor Embedding). The visualization clearly demonstrates the distinguishability of the *leakage state*, indicating the efficiency of the classification. In addition, we find that signals from *Stage I* and *Stage II* can also be well distinguished from each other due to their different signal patterns. After extracting the features, TwinLeak identifies the state of each window using a pre-trained decision tree.

However, the occasional anomalies in Fig 8 tell that we cannot directly trigger an alarm for the leakage once one window of *leakage state* is detected. In practical, we identify a leakage event in the early stage by detecting k successive windows of *leakage state - Stage I* to reduce the false alarms. We empirically set $k = 2$ to achieve a trade-off between the early-stage detection delay and the false positive rate.

B. Leakage Signal Extraction

Now we have detected the leakage, the next target of TwinLeak is to extract the signal segment that contains a complete leakage process for further analysis. This is however a challenging problem because we have to identify the starting and the ending points accurately, or the subsequent jointly estimation process will be error-prone. A naive method is to exploit the result of the state identification process – simply identifying the first leakage-state window as the start point and the last one as the end point. However, this turns out to be brittle because the identification of each individual window is not precise enough. To address this problem, we propose a two-step leakage signal extraction method, which first provides a *coarse segmentation result* based on the state identification results, and then localize the exact starting and ending points on the coarse segment, providing a refined segmentation result.

Step 1: Coarse segmentation. In the coarse segmentation stage, TwinLeak identifies a succession of leakage-state windows and splices the signal contained in these windows to generate a coarse signal segmentation result. The problem we meet here is how to tolerate the errors in state identification. Instead of treating each signal window independently, here we explore its predecessors and successors to eliminate anomalies.

Specifically, given a sequence of labeled signal windows $\{T_j^{(W)}\}$, we calculate a parameter $I_j^{(W)}$ for each window which indicates the ratio of its N contextual windows that are identified as leakage states:

$$I_j^{(W)} = \frac{1}{2\lfloor\frac{N}{2}\rfloor + 1} \sum_{k=j-\lfloor\frac{N}{2}\rfloor}^{j+\lfloor\frac{N}{2}\rfloor} \mathbb{I}(T_k^{(W)}\ is\ leakage) \quad (5)$$

Then, a succession of m signal windows $\{T_j^{(W)}\}_m$ which satisfies (i) all $I_j^{(W)} > 0$ and (ii) its length $m > M$ is extracted as a coarse segmentation of leakage signal. We empirically set $N = 3$ and $M = 4$ in the implementation of TwinLeak.

Step 2: Refined segmentation. Next, we refine the extracted leakage signal by identifying the specific starting point of the leakage event. In Sect. III-B, we observe that the RSSI / phase readings remain stable before the leakage or after the cotton gets saturated. Thus, we can estimate the starting point by calculating the intersection point of the local linear fit of *normal* samples and the local linear fit of *leakage* samples.

However, since our signal source contains four streams: RSSI and phase readings of the twin tags (denoted as $R_i, P_i, i \in \{1, 2\}$, respectively), separately localizing the starting point $sp_{[s]}$ for each stream $s \in S = \{R_1, R_2, P_1, P_2\}$ is error-prone and ambiguous. We choose a joint estimation from the result collection $\mathbf{sp} = \{sp_{[s]} | s \in S\}$ that minimizes the sum of the normalized fitting errors.

VI. Joint Estimation

As we observed in Sect. III-B, there exist temporal and spatial relationships among the signal patterns under different leakage volumes and rates. Therefore, after extracting the signal segments corresponding to the leakage process, TwinLeak then jointly estimates these properties and reconfirms the early-stage detection.

A. Model-Guided Leakage Tracking

To estimate leakage volume and rate, we should first track the leakage process. Therefore, a simple and low-overhead calibration is first performed to get a complete and accurate signal template of the leakage process. Then we track the leakage progress by matching the signal segments extracted fast detection component to the target signal template.

Step 1: Low overhead calibration Tag calibration is conducted only once before the deployment of the tags. During calibration, we manually trigger leakage process on the Twin-Tag sensors, where the leakage rate is stable, and the leakage volume is controlled at $20mL$. The above process is

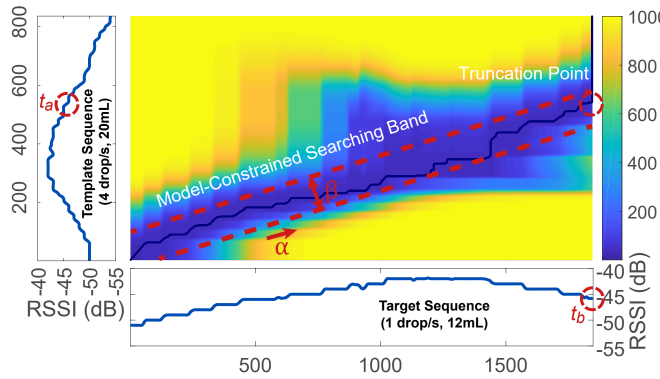


Fig. 9. **Model-Guided Leakage Tracking.**

repeated for 30 times on different sensors and under different leakage rate, during which the reader continuously reads the tag to get the RSSI and phase streams. At last, an nearest-centroid algorithm called DBA [17] is exploited to generate the template based on the collected data.

Step 2: Leakage progress estimation TwinLeak tracks the leakage process using a Dynamic Time Warping (DTW) based search algorithm. A good overview of DTW algorithm can be found in [18]. Broadly speaking, DTW compares two sequence segments, denoted as $\mathbf{S_a} = \{S_a^{(1)}, ..., S_a^{(L_a)}\}$ and $\mathbf{S_b} = \{S_b^{(1)}, ..., S_b^{(L_b)}\}$ by warping (i.e., stretching and squeezing) one of the sequences until an optimal match between them is found. More specifically, the target of DTW algorithm is to find a matching function f_M between $\mathbf{S_a}$ and $\mathbf{S_b}$ so as to minimize $\sum_{i=1}^{L_a}(S_a[i] - S_b[f_M(i)])^2$.

Despite its wide usage, DTW cannot be directly applied to track the leakage process for the following two reasons: (i) The goal of DTW is to minimize the sequence distance after arbitrary warping, however when matching two leakage signals, the warping should be relatively "steady" rather than "arbitrary", due to the relatively stable leakage rate. (ii) The leakage process may be stopped before the cotton is saturated, thus incomplete sequence matching should be considered. Therefore, to address the above problems, TwinLeak proposes a model-guided DTW progress tracking method. As shown in Fig. 9, in this method, we constrain the DTW matching path with 2 parameters α and β, where:

- α guides the direction of searching: since commonly the leakage rate is relatively stable, α – the ratio of the leakage rate of the two leakage signals is relatively stable.
- β bounds the step length of searching: such step-constrained DTW can somehow avoid arbitrary warping and tolerate local jitters.
- Together with α and β, the spatial truncation of "unsaturated" signal can also be achieved in advance to avoid unexpected errors.

We formulate these limitations with the following:

$$|f_M(i_b) - i_a| = |\lfloor\alpha \cdot i_b\rfloor - i_a| \leq \beta \quad (6)$$

The above limitations on the DTW matching path are derived by the leakage signal model observed and analyzed in Sect. III, thus we call it model-guided DTW matching. We note that the value of α can be roughly estimated by comparing the length ratio of *leakage-Stage I* of two leakage sequences, and the value of β can be derived from the statistics of training sequences. An example of our model-guided tracking method is shown in Fig. 9, where the dark blue curve indicates the optimal DTW matching path $\mathbb{M}$ of two leakage signals: the template sequence (left, 4 $drop/s, 20mL$) and the target sequence (bottom, 1 $drop/s, 12mL$). We can observe that $\mathbb{M}$ is restricted in a narrow band, whose direction is determined by $\alpha = \frac{1}{4}$ and the boundary is determined by $\beta = 5$.

B. Property Estimation.

As shown in Fig. 2, the temporal and spatial transformation of the signal patterns are closely related to the leakage rate r and volume v respectively. Given the template sequence $\mathbf{S_a}$ and the target sequence $\mathbf{S_b}$, we conclude the critical observations as (i) $r(\mathbf{S_b})$ determines its temporal scaling of the template $\mathbf{S_a}$ without the shape violation; (ii) $v(\mathbf{S_b})$ does not scale up or down $\mathbf{S_a}$ but spatially truncates it by entering the stable state in advance. Thus, based on these observations, $r(\mathbf{S_b})$ and $v(\mathbf{S_b})$ can be estimated by inferring the temporal scaling factor and spatial truncation ratio.

In TwinLeak, we jointly estimate the probability of the leakage event and infer the leakage rate and volume by inspecting the model-guided leakage tracking process. As shown in Fig. 9, the truncation point $t = (t_a, t_b)$ is detected through inversely traversing by maximizing the similarity between the truncated subsequences $\hat{\mathbf{S}}_\mathbf{a} = \mathbf{f}_\mathbb{M}(\mathbf{S_a}[1:t_a])$ and $\hat{\mathbf{S}}_\mathbf{b} = \mathbf{f}_\mathbb{M}(\mathbf{S_b}[1:t_b])$. Then, we can jointly estimate the *average matched distance* $d(\mathbf{S_b}) = \frac{1}{t}\sum_{i=1}^{t}(\hat{\mathbf{S}}_\mathbf{a}[i] - \hat{\mathbf{S}}_\mathbf{b}[i])^2$, the leakage rate $r(\mathbf{S_b}) = r(\mathbf{S_a}) \cdot \frac{t_a}{t_b}$ and volume $v(\mathbf{S_b}) = v(\mathbf{S_a}) \cdot \frac{t_a}{L_a}$. To estimate the probability of the leakage event, we calculate the probability that $d(\mathbf{S_b})$ belongs to a pre-derived distribution of the average matched distances of the training set. This probability could be further utilized as an reconfirmation of the early-stage detection made by fast detection component.

VII. Implementation and Evaluation

In this section, we first detail the implementation of TwinLeak (Sect. VII-A) and then introduce the evaluation settings and the metrics (Sect. VII-B). The experiment results of some critical micro-benchmarks and the detection precision is shown in the subsequent subsections.

A. Implementation

The implementation of TwinLeak mainly consists of the hardware part and the software part. First, we handcraft dozens of the Twin-Tag. Two important parameters of the Twin-Tag are: (1) the size of the absorbent cotton is around $10cm \times 2cm \times 0.4cm$, which can absorb about $20mL$ liquid before getting saturated; (2) the distance between the twin tags is around $1mm$, which can result in about $5dB$ RSSI gap. Apart from the sensor, an ImpinJ Speedway R420 RFID reader with a Laird circular polarized antenna is deployed to provide continuous waves and receive the backscattered signals from Twin-Tag. Fig. 10 demonstrates the real-world deployment of TwinLeak in our industrial IoT project Pavatar for the power industry [19].

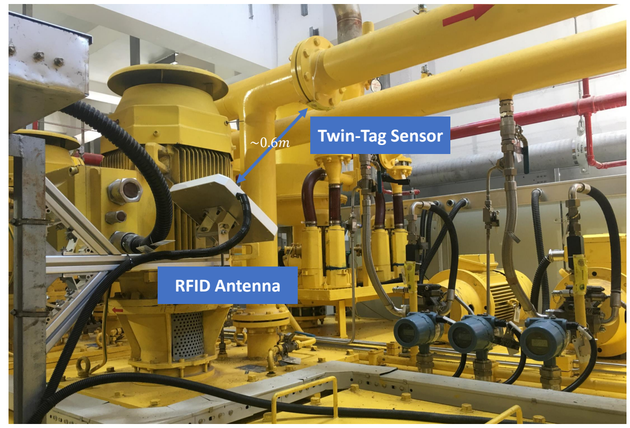


Fig. 10. **Real-World Deployment of TwinLeak.**

The software part has 3 main components: (1) a Websocket data provider (implemented in Java) that continuously acquires RSSI and phase readings through ImpinJ's API; (2) an offline trainer (implemented in MATLAB) that generates the general signal templates; (3) an online detector (implemented in Python) performs the liquid detection task with the processing components described in Sect. V and VI.

B. Methodologies and Settings

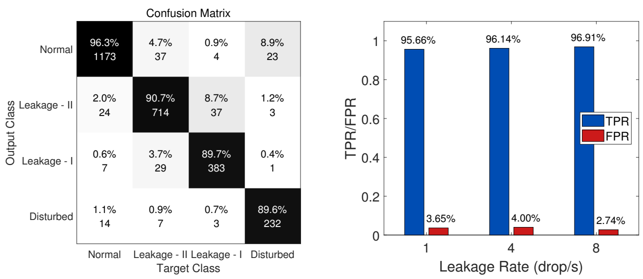
Our evaluation mainly focuses on the problem of detecting whether the leakage event occurs or not in given signal sequences. TPR (True Positive Rate) evaluates the capability of TwinLeak to recognize the leakage event, and FPR (False Positive Rate) evaluates the frequency of false alarm under the multi-path interferences. We also focus on the impacts of the critical properties (e.g. rate and volume) of the leakage event on the detection accuracy.

The data for training the state identifier and extracting the signal template is collected in various scenarios, e.g. different leakage rates and volumes as well as unrestricted reader-tag distances. The signals disturbed by the external multi-path interferences are also utilized for the offline classifier training. Furthermore, we separately collect an evaluation dataset, in which the leakage rates $(1, 4, 8drop/s)$ and volumes $(10, 15, 20mL)$ are the main variables, and reader-tag distance is fixed to $60cm$ since we measure that such distance won't change the signal pattern in Sect. III-B. We collect 10 samples for each combination of these two parameters.

C. Micro-benchmarks

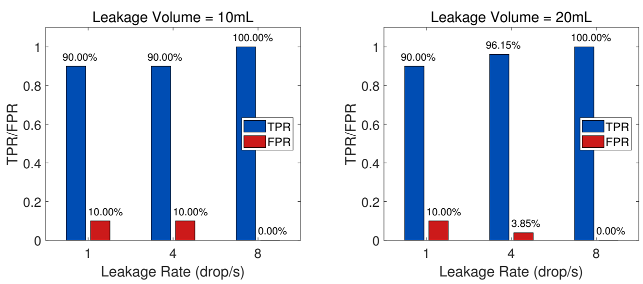
First, we evaluate some micro-benchmarks of the core components in TwinLeak.

(1) State Identification. As the first step of TwinLeak, *state identification* gives the system a preliminary view of the signal state. Fig. 11(a) plots the classification confusion matrix. We can observe that, although we only use a simple but fast



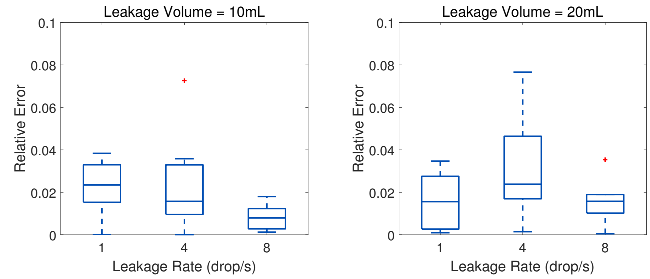
(a) Confusion Matrix (b) Impact of Leakge Rate

Fig. 11. **The evaluation of State Identification.**



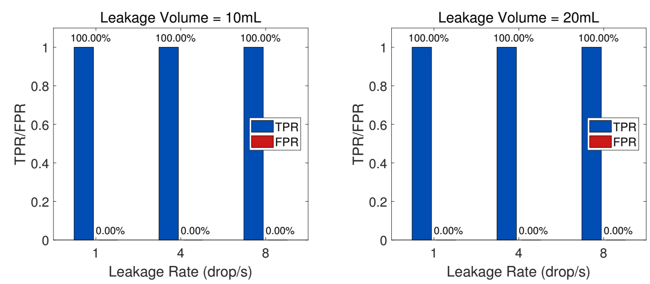
(a) Coarse Seg. Acc. ($v = 10mL$) (b) Coarse Seg. Acc. ($v = 20mL$)

Fig. 12. **The evaluation of Coarse Segmentation.**



(a) Local. Error ($v = 10mL$) (b) Local. Error ($v = 20mL$)

Fig. 13. **The evaluation of Refined Segmentation.**



(a) Temp. Match. Acc. ($v = 10mL$) (b) Temp. Match. Acc. ($v = 20mL$)

Fig. 14. **The evaluation of Model-Guided Leakage Tracking.**

decision-tree classifier, three signal states can be well distinguished from each other due to the good patterns provided by Twin-Tag. Although "inner" misclassification of *Stage I* and *II* of the leakage state is relative high, the TPR of overall leakage state reaches 98.7%. Moreover, the false alarm induced by the interferences is acceptable although it might be misclassified as the normal state. Since *state identification* adopts a fixed time-window of length $10s$ for feature extraction, we evaluates the impact of the rate of the signal variation by controlling the leakage rates. Fig. 11(b) shows that the FPR and TPR are almost unaffected by the leakage rate, which mainly results from the good generalizability of Twin-Tag's signal patterns.

(2) Leakage Signal Extraction. The two-step signal extraction method, which extracts the signals during the leakage events from coarse to refined, is a critical pre-processing further false-alarm refusal and property estimation.

We denote the true positive of *coarse segmentation* as the result that it generates a good proposal containing an ground truth leakage signal. We can tell from Fig. 12 that an incomplete coarse segmentation is more likely to be proposed due to the misclassification of data slices during the time when the signal variation is relatively gentle. e.g. $r = 1\ drop/s$. We will address later that these incomplete segmentations won't affect the final classification but might degrade the property estimation.

Besides, we measure the relative errors of *refined segmentation* on those good sequences proposed by the last step. Fig. 13 shows that low relative errors (the absolute localization errors divided by the total length) around 1.5% can be obtained by the algorithm. The rate of signal variation mainly affects the performance, because the deviations get more distinct if the signals vary more sharply.

D. Leakage Detection

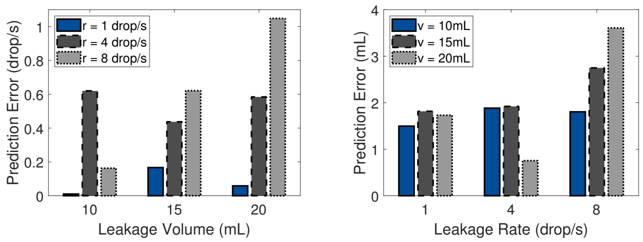
Second, we evaluate the main task of the leakage detection by examining the final detection accuracy and the prediction errors of property estimation.

(1) Detection Precision. We match all the segmented testing data ($3 \times 3 \times 10 = 90$ samples), provided by our fast detection module, to the pre-learned signal template with the proposed model-guided leakage tracking algorithm. The average matched distance d is derived from the matching process, which solely acts as the final indicator for the binary classification. Fig. 14 demonstrates that TwinLeak can achieve a relatively high TPR and low FPR regardless of the leakage rates and volumes. Although the segmentation might be incomplete, the optimal average matched distance along the restricted searching band can well describe the similarity between the template and the target sequences.

(2) Estimation Error rate. Then, given the leakage rate $r_{\overline{T}} = 4\ drop/s$ and volume $v_{\overline{T}} = 20mL$ (Saturated) of the template, we evaluate the prediction error of the leakage rate and volume. As shown in Fig. 15, we can tell the absolute prediction errors for both of the leakage rate and volume are very small and not sensitive to the changes of the other property. Moreover, we calculate the average relative prediction error as $\tilde{e}_r = 14.7\%$ and $\tilde{e}_v = 15.3\%$.

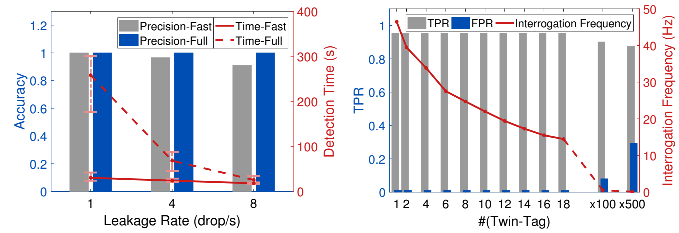
E. Application-Driven Metrics

(1) Timeliness V.S. Accuracy. To satisfy the timeliness requirement, TwinLeak can directly raise an alarm when k successive windows are determined as the *leakage state - Stage I* (denoted as *TwinLeak-Fast*). Compared to the full procedure of TwinLeak (denoted as *TwinLeak-Full*), TwinLeak-Fast can improve much timeliness but might sacrifice some



(a) Leakage Rate Estimation (b) Leakage Volume Estimation

Fig. 15. **The evaluation of Property Estimation.**



(a) Timeliness V.S. Accuracy (b) Multi-Point Surveillance

Fig. 16. **The evaluation of Property Estimation.**

performance. Fig. 16(a) shows the trade-off between timeliness and accuracy. We can see that the slower the leakage process, the greater the gain of TwinLeak-Fast (e.g. about 8.5X when $r = 1\ drop/s$). We find that TwinLeak can make early leakage alarms when only about $4.6mL$ liquid leaks in average.

(2) Multi-Point Surveillance. Another critical application-driven evaluation is the capability of monitoring multiple potential leakage points, e.g. numerous and dense flanges shown in Fig 1. It is known that the interrogation frequency of a RFID tag will degrade as the increase of the number of its densely-deployed neighboring tags due to the channel collision [20]. However, TwinLeak are competent for the multi-point surveillance task, because its tendency features are very resilient to the degradation of the interrogation frequency. As shown in Fig. 16(b), TwinLeak can achieve an overall TPR higher than 97.2% and an overall FPR lower than 0.5%, and the performance has barely declined, even in the case of very intensive deployments (18 Twin-Tags, $\sim 3X$ lower interrogation frequency). We further simulate a stronger degradation in the interrogation frequency by random down-sampling and find that when it declines more than $100X$ our starts to become unavailable.

VIII. Conclusion

In this work we present TwinLeak for real-time and accurate leakage detection using COTS RFID tags and readers. A key innovation is to detect leakage and estimate leakage rate and volume based on the changes in the strength of coupling effect between tags. TwinLeak achieves a TPR higher than 97.2% and keeps the FPR below 0.5%, with the early alarms after only about $4.6mL$ liquid leaks, providing the necessary precision and delay for many applications. Moreover, TwinLeak can simultaneously estimate the leakage properties e.g. rate and volume, and keep the relative prediction error around 10%. TwinLeak not only has been tested and used in practical applications, but also will open up a wide range of exciting research opportunities.

IX. Acknowledgment

This work is supported in part 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. We thank all the anonymous reviewers for their valuable comments and helpful suggestions.

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] 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.

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

[5] 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.

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

[7] Ashutosh Dhekne, Mahanth Gowda, Yixuan Zhao, et al. Liquid: A wireless liquid identifier. In *Proceedings of ACM Mobisys*, 2018.

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

[9] Jinlan Gao, Johan Sidén, Hans-Erik Nilsson, et al. Printed humidity sensor with memory functionality for passive rfid tags. *IEEE Sensors Journal*, 2013.

[10] Ivan Stoianov, Lama Nachman, Sam Madden, et al. Pipeneta wireless sensor network for pipeline monitoring. In *Proceedings of IEEE/ACM IPSN*, 2007.

[11] Ali M. Sadeghioon, Nicole Metje, David N. Chapman, et al. Smartpipes: smart wireless sensor networks for leak detection in water pipelines. *Journal of sensor and Actuator Networks*, 2014.

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

[13] Sabina Manzari and Gaetano Marrocco. Modeling and applications of a chemical-loaded uhf rfid sensing antenna with tuning capability. *IEEE Transactions on Antennas and Propagation*, 2014.

[14] Jinsong Han, Chen Qian, Xing Wang, et al. Twins: Device-free object tracking using passive tags. *IEEE/ACM Transactions on Networking*, 2016.

[15] Ge Wang, Haofan Cai, Chen Qian, et al. Towards replay-resilient rfid authentication. In *Proceedings of ACM MobiCom*, 2018.

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

[17] François Petitjean, Alain Ketterlin, and Pierre Gançarski. A global averaging method for dynamic time warping, with applications to clustering. *Elsevier Pattern Recognition*, 2011.

[18] Young-Seon Jeong, Myong K Jeong, and Olufemi A. Omitaomu. Weighted dynamic time warping for time series classification. *Elsevier Pattern Recognition*, 2011.

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

[20] Qiongzheng Lin, Lei Yang, Huanyu Jia, et al. Revisiting reading rate with mobility: Rate-adaptive reading in cots rfid systems. In *Proceedings of ACM CoNext*, 2017.