Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 9, No. 4, 2021, 425-438.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

Water Meter Replacement Recommendation for Municipal Water Distribution Networks using Ensemble Outlier Detection Methods

Fatemeh Kaveh-Yazdy and Sajjad Zarifzadeh*

Computer Engineering Department, Yazd University, Yazd, Iran.

Article Info

Article History:
Received 01 April 2021
Revised 07 May 2021
Accepted 26 June 2021

DOI:10.22044/JADM.2021.10672.2202

Keywords:
Water Metering, Apparent Loss, Failure Detection, Outlier Detection, N-gram, Time Series Analysis.

**Corresponding author: szarifzadeh@yazd.ac.ir (S. Zarifzadeh).*

Abstract

Due to their structure and usage conditions, water meters face degradation, breaking, freezing, and leakage problems. There are various studies intended to determine the appropriate time to replace the degraded water meters. Earlier studies have used several features such as meteorological parameters, usage conditions, water network pressure, and structure of meters in order to detect the failed water meters. This article proposes a recommendation framework that uses the registered water consumption values as the input data and provides the meter replacement recommendations. This framework takes the time series of registered consumption values and pre-processes them in two rounds in order to extract the effective features. Then the multiple un-/semi-supervised outlier detection methods are applied to the processed data and assigns outlier/normal labels to them. At the final stage, a hypergraph-based ensemble method receives the labels and combines them to discover the suitable label. Due to the unavailability of the ground truth labeled data for meter replacement, we compare our method with respect to its FPR and two internal metrics: Dunn index and Davies-Bouldin index. The results of our comparative experiments show that the proposed framework detects more compact clusters with a smaller variance.

1. Introduction

Water loss is a crucial issue in the water distribution networks. The International Water Association (IWA) defines the water losses as the difference between the volume of input water and authorized consumption [1], and divides the water losses into two categories with respect to the reason behind the loss, as follows [2]:

- Real losses
- Apparent losses

The annual real loss value is the total amount of water losses through fatalities in the distribution network, and includes the leakage, bursts, overflows on mains, and in-service losses. On the other side, the annual apparent loss addresses the volume of water losses due to unauthorized consumption and measurement inaccuracies. In contrast to the real losses that happen prior to the consumers' meters, the apparent losses occur after the meters.

In addition to the environmental value of lost water, the losses reduce the revenue of the water supply industry. Arregui *et al.* [3] have indicated that the total volume of apparent losses is less than the volume of real losses. However, the financial loss caused by the apparent losses is more than that resulting from the real losses. Every cubic meter of real loss costs as much as produced water but the cost of same volume of apparent loss equals to the retail water price paid by the customers [3].

Clearly, the expense of apparent losses for water plants is broken down into two parts: (1) cost of water thefts and (2) cost of water inaccurately measured by faulty meters. Thus, two major problems could be defined in this area. The first problem addresses detecting unauthorized water consumption (a.k.a. water theft), and the second problem targets determining the faulty water meters. These two problems are twisted together, while the registered values of inaccurate meters are not reliable for unauthorized consumption detection. Subsequently, determining failures in water meters is a prerequisite for substituting/calibrating them in order to keep the measurement error in a tolerable range.

Typically, a water meter is replaced due to the measurement inaccuracy resulting from degradation and aging or unexpected failures such as burst and freezing [4]. Even though the freezing events can be predicted with respect to the meteorological parameters, the bursts with exogenous causes (e.g., earthquakes and accidents) are not predictable. On the other hand, water meters are gradually degraded, and thus register the values erroneously [3].

Arregui *et al*. [3] have noted that only a few countries had developed standards in order to determine the water meter degradation under various meteorological conditions. Instead, the water supply companies aim at recognizing the error in water meters using data analysis techniques. The faulty meter detection methods can be divided into three groups, as follow:

- Methods that use mechanical properties of meters;
- Methods that utilize mixed-type of key factors affected by aging;
- Methods that adopt consumption records and meters' properties in order to determine the anomalous over-/under-registration values.

The first group of methods uses the installment conditions and metallurgical properties of the components and considers the meteorological variables of the installation region to predict the longest period during that the meter meets the measurement requirements.

As one of the oldest studies, Orr *et al*. [5] have compared the mechanical water meters under various operational conditions in order to determine the longest period of time for their active usage. They developed an optimized replacement plan to upgrade the mechanical water meters to the magnetic-drive ones at the latest possible time. Stoker *et al*. [6] have shown that some water meters meeting the American Water Works Association (AWWA) standards on the installation time might fail to meet these standards in a short period of time after installation, while some others could preserve the standards even after two or three decades that is beyond their expected service lifetime. The service age had been the key factor in making the replacement decisions for a long time; however, the results of this study indicated that the meters' degradation also depended on other features like water velocity, water quality, wear, and throughput value as well as service age; subsequently, the estimated accuracy degradation is utilized as the key factor for devising a proper policy for the water meter replacement.

Clincui and Clincui [7] have investigated the cold-water single-jet water meters. In contrast with the conventional method used in earlier studies to determine the overall reliability of a water meter, they determined the reliability of its functional components. The reliability values of the components of a meter are combined through a numerical model with respect to the links between the components. The results of the study showed that the predicted reliability of the meters was lower than their operational reliability due to the degradation resulting from the usage conditions. They suggested that the replacement plans took into account the properties of the components separately, and integrated them into a model.

As one of the earliest studies, Arregui *et al*. [3] have selected two groups of single-jet water meters as the most frequently used ones in the residential water networks, and evaluated their initial and after installation error values. They presented a list of tests in order to determine the after-installation error values regarding the meteorological parameters and the in-service conditions. The collected observations were employed in order to develop three non-linear degradation models that were able to predict the degradation level of water meters and recommend the replacement cases. Although there is not any report upon an integrated system used to predict the degradation, Alegre *et al*. [1] have compiled the IWA Performance Indicator System report, providing a standard benchmark and performance assessment framework.

The second group of methods extracts the mixed-type key factors and utilizes them in the performance prediction systems in order to detect the degraded and faulty water meters. The parameter set includes the mechanical properties, position, metallurgical properties of meter, meter age, pipeline size, network pressure, flow velocity, water quality, and total metered water. Consolidating the data of diverse types with various range of values is challenging, and requires expert involvement, which is distinct for this class of methods.

Arregui *et al*. [8] have investigated the accuracy of five groups of industrial and domestic water meters under different real-world and laboratory-simulated situations in order to evaluate the impact of a wide range of factors such as the mounting position, velocity profile distortions, fatigue test results, depositions, partial blockage of the strainer, incorrect water meter sizing, and water consumption patterns. Their tests showed that not only different classes of water meters but also diverse sample meters present very different erroneous behaviors. These findings were reported again in the later studies considering the commercial and residential water meters [9].

Fontanazza *et al*. [10] have proposed an indicator aggregating the meter's age, total metered water volume, and network pressure. They applied multiple weighting schemas and utilized a Monte Carlo simulator in order to evaluate the robustness and reliability of these factors in meter replacement planning. Yazdandoost and Izadi [11] and Cardoso *et al*. [12] have aggregated the mechanical, functional, and network parameters in the asset management models in order to measure the performance of the meters and suggest the replaceable meters.

The impact of the flow of water on the accuracy of the meter is inevitable, and this flow is a function of the difference between the network and the consumers' pipeline pressures. Fontanazza *et al*. [13] have used the mixed-type data collected through a monitoring campaign and designed a numerical model to estimate the apparent losses based on the volume of water flow. In their model, they took into consideration the possibility of using a reserve tank in the consumer's pipeline network. This study suggests a replacement strategy based on the time window of the maximum losses for each meter. Weller *et al*. [14] have studied the impact of surge flows as well as other types of features on the accuracy of different classes of residential water meters. They defined the surge water flows as flows that have a higher rate than that defined by the AWWA standards for the residential meters. In this study, the accuracy of water meters was evaluated before, after, and during the surge. Their results indicated that the mechanical meters were highly sensitive to the surge flows, and they were more likely to present under-register values for low velocity flows after being damaged by a surge flow.

The electronic meters are unlikely to be damaged by the surge flows. However, their accuracies are decreased after very high rates of surge flows. The authors ended up their report with a recommendation plan that suggested appropriate classes of water meters for a network facing the possibility of surge water flows.

The third group of methods analyzes the consumption behavior of the consumers, and tries to find the anomalous patterns resulting from the leakage, bursts, meter degradation, and water theft. This group detects the behaviors relevant to the apparent losses, and reports them for further investigations. We should note that most of the conducted research works have used the mechanical and functional characteristics of the meters in order to decide about their degradation and recommend replacement; however, there are a few studies that use other kinds of anomalies in the water networks. As an example, Kanyama *et al*. [15] have utilized the collected measurements of smart wireless water meters to detect the anomalous registering behaviors in order to find the records that may be injected by attacking the wireless network. Andrysiak *et al*. [16] have proposed an advanced time series anomaly detection for detecting the outlier records in smart wireless water networks. They focused on the anomalies found in the received measurements of water meters. Moahloli *et al*. [17] have analyzed the consumption time series of the consumers showing that the total registered volume has a positive relation with the jet meters' error. They found that the erroneous values registered by the meters move from the under-registration to the over-registration values as the total volume of water increases, which arises from impeller degradation.

In this work, we target the problem of recognizing the failed water meters by employing the easy-to-access water consumption information in our selected case city, which is the city of Yazd in Iran. Since the water meters' properties have not been collected and stored in an integrated database, the proposed method relies only on the water consumption records. To the best of our knowledge, the proposed framework is the first instance system that uses only the consumption data and does not take into account the meters' properties.

The previous studies have shown that the amount of water and energy consumption for the consumers reflect a periodic behavior, which primarily depends on the weather and history of consumption [18]–[22]. Our idea is that the collected measurements of the normal consumers must present periodic patterns, and consequently, the aperiodic patterns detected in the registered consumption values can be considered as the anomalous behaviors, and an evidence of meter failure.

The proposed system is an empirical framework designed to deal with the following challenges:

- Un-labeled records

  The history of meter replacement is available in the database of the Yazd Water and Waste-water Company (YWWC) but the reason behind the replacement and the results of the control tests that have been done to determine whether the meter is faulty are not stored.

- Water meters' data lack

  The installation date, producer, and mechanical properties of the meters are not stored.

- In-available demographic data

  The tracking of consumption behaviors without demographic data of households is challenging, but YWWC does not provide the mentioned data.

- Missing values

  More than 30% of the consumption records include at least one missing value.

- Mixed sampling rate

  The registered values of meters are collected manually, and in a few cases, based on the consumers' requests. Accordingly, the days between two registered values can be varied from 5 to 281 days.

- Lack of expert knowledge about the in-service meter replacement plan

  At the present time, the faulty meters are detected by a self-report and a decision support system that is not deterministic and well-documented.

Due to the lack of data and the presence of noises, we cannot utilize the conventional methods, and we have to propose a novel framework in order to overcome the mentioned challenges and detect the faulty meters with a tolerable error.

Our major contributions are as follows. First, we present a practical solution for the large-scale meter replacement with a low false-positive rate. Secondly, while the degraded replaced meters tend to be mixed with the meters that are replaced due to freezing and breaking, the proposed method is robust against such replacements. Finally, our method benefits from an ensemble of semi-supervised outlier detection methods in order to distinguish the meters presenting different kinds of abnormal measurement behaviors.

The rest of this paper is organized as follows. Section 2 defines the water meter replacement problem. Section 3 covers the data collection, pre-processing, and feature selection steps. The details of the proposed framework and its stages are presented in Section 4. The experiments and their results are presented and discussed in Section 5. Finally, we conclude the paper in Section 6.

2. Problem Definition

In Iran, as our testbed, customer's self-report, the report by the person who collects the meter values, and a rule-based decision support system (DSS) provide the signal for water meter replacement. When at least one signal triggers, a technician visits the meter and decides whether it must be replaced or not. This approach suffers from the noise that comes from human errors. Regardless of the reporting error, the rules that are embedded in DSS are developed by human experts, which makes them vulnerable to bias. Indeed, the DSS system can only detect a specific group of degradation patterns represented by its rules, and the unseen patterns belonging to the reported meters are not taken into account for future replacement.

Intuitively, the time series of values registered by the in-services meters that are not degraded have uniform behaviors. Consequently, a recommender system can use the history of registered values of meters and apply the outlier detection methods to extract the replacement candidates. Finally, a technical team visits the site and decides upon replacement.

While the definitive reason behind the replacement is not registered in logs of YWWC, the performance of the proposed system is measured according to the overlap of the candidate cases with the replaced meters. In order to address the meters that have been replaced, we use the term "suspicious", pointing to the uncertainty behind labeling the meter as faulty.

Trivially, the majority of patterns among the class of replaced meters belongs to the meters that are suspected to be faulty (as they are probably replaced due to a problem in their operation or measurement), and the minority of meters found in the group of in-service meters are the defective meters that have not been detected yet. In fact, the water replacement recommendation can be defined as clustering the meters with respect to their registered values.

2. Data

The main data used in this work was provided by YWWC. In the following sub-sections, we briefly describe the dataset and also the preparation policies we designed for this problem.

3.1. Raw Billing Record Extraction

Our dataset included 12,250,988 billing records of 193,573 unique residential consumers belonging to the time period from 21/3/2001 to 20/3/2018. Each billing record had 46 data fields included the consumer's identity number, period start date, period end date, meter registered value on the start date, meter registered value on the end date, consumption, number of days, in-water diameter, previous debt, in-water pressure, tariff code, bill value, etc. Other fields involve the financial and accounting information that is not applicable to our problem.

We should note that the time windows of billing records are varied between 5 to 281 days, based upon the time between the inspector visits and the requested bills. All the water meters used in the Yazd water network belong to the same mechanical class and are produced by at least nine different producers. Unfortunately, the degradation, register error, in-service age, and class of water meters are not provided by YWWC. The only available information about water meters that can be extracted from the billing records is the replacement dates.

According to the replacement records, there are 38715 unique consumers who have at least one replacement record, and the whole data includes 44706 meter replacement cases. These records are labeled as suspicious records because the reason behind the replacement is not recorded, and none of these meters can indeed be labeled as degraded or fatal. In order to light a shed on this fact, suppose that a meter works normally, but it is broken due to winter freezing and is then replaced. The mentioned meters might register the consumption values correctly before the freezing event. Thus its historical records do not probably show any anomalous pattern.

The only possible and reliable solution is to learn the consumption behavior of the consumers who have no replaced-meter record in their history and detect the anomalous behavior that happened in degradation age by comparing these two behavioral patterns. For this purpose, we extracted two sets of consumers from our data and then retrieved their billing records: (1) the set of all consumers experiencing a replaced meter, called the suspicious consumers, and (2) the set of randomly selected consumers without any meter replacement, called the clean consumers. For the suspicious consumers, we selected enough consumption records before the replacement date. The start points of the time windows for clean consumers were randomly chosen. The retrieved record set for 38715 suspicious consumers included 301940 billing records, which we call suspicious record set, and that of the set for 68318 clean consumers containing 1512767 records, named the clean record set.

4. Proposed Framework

In this section, we present an outline of the proposed framework and then review each step briefly. The proposed framework intends to find the water meters that are likely to be degraded with respect to their recent registered consumption values. Figure 1 shows the structure of the proposed framework. Each transaction in this system is initiated by an investigation request issued by one of the YWWC customer service desk investigators who requests the recommended action upon replacing the water meter of a specified consumer.

The request is received by the Yazd municipal water billing system. Then the recent billing records of the targeted consumer are attached to the request and routed to the proposed recommender framework. The raw data of records is pre-processed through several processes including feature generation, feature selection, normalization, and symbolization.

The next step would be to apply the trained outlier detection models to label the record. Finally, an ensemble classifier decides about the record and assigns a definitive label. The label is returned back to the Yazd municipal water billing system for the announcement. At the end, if the suggested label is "anomalous", then the meter replacement decision is recommended to the technical team.

4.1. Data Preparation: First Round

In this section, we outline the first round of data preparation, which includes feature generation, data normalization, and feature selection. This step is fed with raw billing records, and converts the records into the z-normalized series. The impacts of the features on the distinguishing clean records (belonging to the working meters) from the suspicious records (belonging to the replaced meters) are measured, and the effective features are selected.

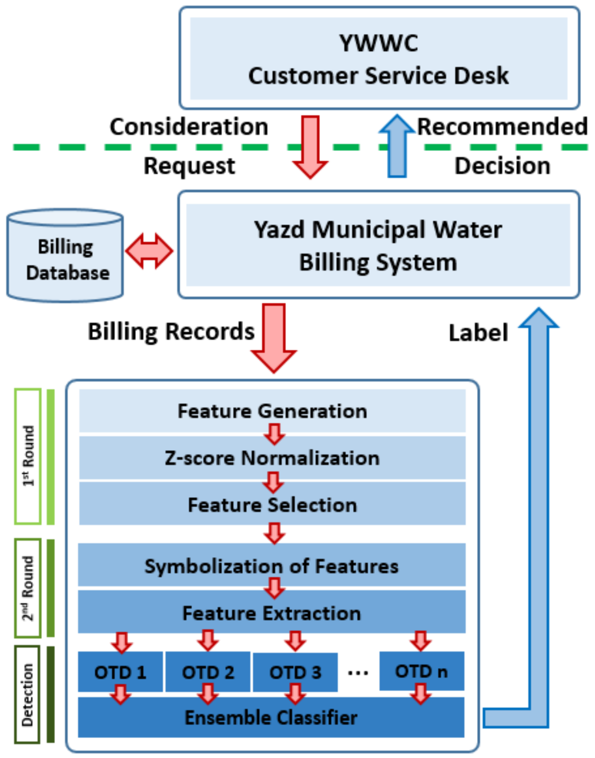

Figure 1: A schematic view of the proposed water meter replacement recommender system.

4.1.1. Feature Generation

The past research studies have considered both the pressure and the consumption history as the important data fields in the aforementioned billing database. In addition, the experts at YWWC suggest the previous debt as a potential indicator of meter degradation. This is due to the fact that the consumers who are doubtful about the accuracy of their water meter often refuse to pay their bills before the meter check or replacement. Therefore, the pressure, consumption history, and previous debt fields are considered as potential features.

The YWWC's technical team suggests that the selected records cover a period of six to fourteen months ([180, 425] days). Unfortunately, the time windows of the billing records are not equal; thus, we divide the consumption value of each billing record by the number of days in that record in order to compute the consumption per day.

4.1.2. Z-score Normalization

During our investigations, we found that the values of the previous debt and daily consumption were significantly different between the consumers, and hence, we were required to normalize them first. The z-score normalization appeared to be a proper normalization method to prepare the raw records. The Z-score normalization uses the mean and standard deviation (STD) of values in order to calculate the normalized value of v_i as follows:

$$v' = \frac{v_i - \bar{v}}{\sigma_v} \tag{1}$$

where $\bar{v}$ and σ_v are the mean and standard deviation of values, respectively. The interval-scaled pressure values are in the range of [1.52, 3.55]. Due to the invariant distribution and a limited range of values, the pressure values do not require Z-score normalization. Up to this point, the generated and normalized feature set includes the Z-score normalized daily consumption, Z-score normalized previous debt, and pressure.

4.1.3. Feature Selection

In this step, a discrimination analysis [23] is used in order to select the features that are able to distinguish the clean consumers from the suspicious ones. The dissimilarity of the distribution of feature values in both sets must be able to distinguish them; therefore, we select the Jensen–Shannon Divergence (JSD) in order to discriminate the sets against each feature (similar to [24], [25]). The similarity of the probability distribution of features in both record sets is computed using a 50-bin probability distribution generated for each feature.

JSD is symmetric smoothed version of the Kullback–Leibler Divergence (KLD), which is able to indicate the similarity of the probability distribution. The JSD of two discrete probability distributions can be measured by computing the following equation:

$$\begin{aligned} SD(P \| Q) &= \frac{1}{2} D(P \| M) + \frac{1}{2} D(Q \| M) \\ M &= \frac{1}{2}(P+Q) \end{aligned} \tag{2}$$

where the parameters P and Q are the discrete probability distributions, M is the average of probability distributions, and $D(P \| M)$ is the KLD of P and M distributions, i.e.:

$$D(P \| M) = \sum_{x \in X} P(x) \log\left(\frac{P(x)}{M(x)}\right) \tag{3}$$

where $P(x)$ and $M(x)$ are the probability values of the discrete random variable x. JSD of daily consumption values, previous debt, and pressure are 0.24, 0.12, and 0.066, respectively. Among the mentioned features, the consumer values are more discriminant than the others and can be utilized in our future analysis.

Accordingly, the pre-processed and selected data records for replacement prediction are time series of normalized daily water consumption values belonging to the time windows of [180, 425] days.

4.2. Data Preparation: Second Round

This section presents the SAX symbolization that is employed to overcome the mixed-sampling rate challenge. Then we benefit from the mean and standard deviation values to highlight the outliers.

4.2.1. Symbolization of Features

While the time window of meter readings is variable and the time series are arbitrarily sampled, the fixed-frequency time series analysis methods are not applicable to the generated consumption time series. Bagattini *et al*. [26] have studied the sampling problem, showing that the conventional symbolic aggregate approximation (SAX) representation [27] is an appropriate representation for the time series with missing values and also unequal time windows. Moreover, they proposed a novel SAX-based schema, which enabled them to reveal the class-distinctive discrete event sub-sequences. Therefore, we use SAX to represent the normalized daily water consumption time series. The SAX method provides a standard breakpoints schema, which partitions the domain of values into the equally-sized partitions with respect to the area under the Gaussian curve.

The next step is to assign the alphabets to the partitioned time series. In this research work, we partition the time series into ten levels applying the Lin *et al*. [28] breakpoint lookup table and assign values from 1 to 10 to them. Subsequently, the daily water consumption series are converted to sequences of 1 to 10 discrete values.

4.2.2. Feature Extraction

As we noted earlier, the problem of detecting faulty meters can be defined as finding the record sets that behave against the prevalent patterns of the working meters. Therefore, we can re-formalize the problem of finding replaceable water meters to an outlier detection problem.

The trends of daily consumption values can describe the normal as well as the exceptional consumption behaviors. Subsequently, if we cluster them into the normal and outlier groups, then we are able to detect the series that are suspicious to belong to the faulty meters.

Leys *et al*. [29] have compared two approaches with respect to their achieved performance in the problems with changes in the trends of values. The first approach uses the mean and standard deviation, and the second one employs the median and absolute deviation from the median. They reported that the first approach was highly sensitive to the outliers, and the second method was robust to them. Accordingly, the first approach indicates the outliers better than the second one. Thus we extract the new features, i.e. the mean and standard deviation values of the symbolized series, and substitute them with the series. In order to explain how the changes in the mean and standard deviation uncover the outliers, we add the results of our data investigation step in Appendix A.

4.3. Outlier Detection

The definition of our problem comes under the umbrella of semi-supervised outlier detection problems [30]. In the semi-supervised outlier detection problems, the learning model is trained by the normal samples to assign the normal/outlier labels to the test set members.

In a similar manner, we use the $\langle mean, STD \rangle$ pairs of clean records to train a semi-supervised outlier detection model, and then employ this model to label all pairs generated from suspicious records.

Figure 2 shows two bivariate distribution plots where each plot demonstrates the bivariate distribution of the mean and standard deviation values in its middle area and the univariate distribution of variables on the mirror sides of the axes.

As it can be seen, the bivariate and univariate distribution of values in the clean and suspicious sets are different; however, they share similar patterns in the area restricted by the constraints $5 \leq mean \leq 6$ and $2.8 \leq STD \leq 3.5$. We should note that the similarity between the overall picture of the mentioned area does not occur by chance but rather, it occurs as the outcome of data mixture.

In this step, the training set is made up of the mean and standard deviation values of the clean records, and five semi-supervised outlier detection methods are applied to it. Then the fitted models are utilized to assign binary labels, which are True for the outliers and False for the normal records.

Among the un/semi-supervised methods introduced by Goldstein and Uchida [30], we select the K-Nearest Neighbor (KNN), Histogram-Based Outlier Scoring (HBOS), Isolation Forest Outlier Detection (IOD), CLuster-Based Outlier Factor (CLBOF), and Local Outlier Factor (LOF). These methods are selected from the local and global nearest neighbor, global clustering-based, and statistical classes of the outlier detection methods.

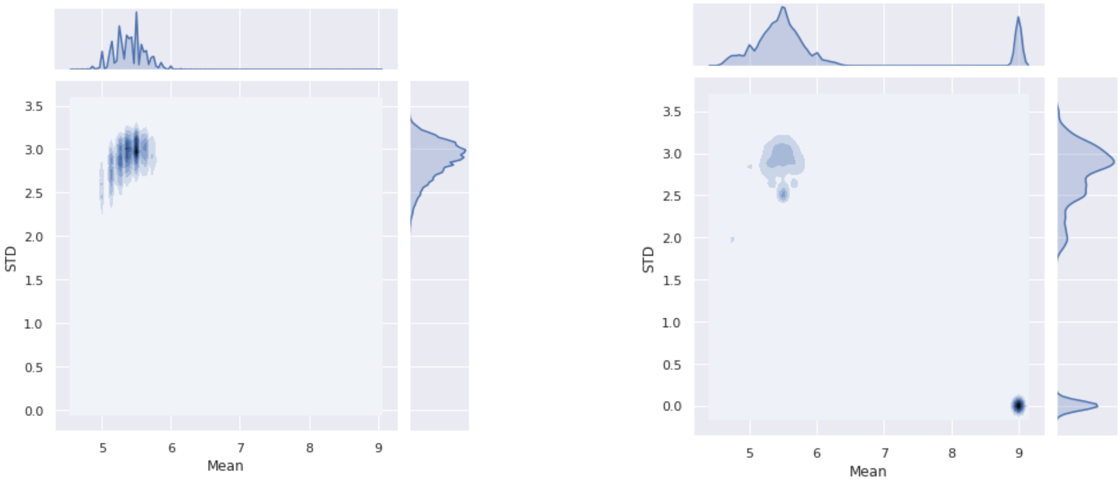

Figure 2: Joint distribution of mean and standard deviation values in both the clean (right) and suspicious (left) record sets.

The primary motivation behind these choices is their low computational complexity and generality to cover various kinds of outliers.

4.4. Ensemble Outlier Detector

Although the labels assigned by the outlier detection algorithms are applicable to the samples, we decided to combine the outlier detector algorithms to achieve more accurate labels [23]. The simplest method for combining the results in ensemble methods is majority voting, in which the ensemble label equals the label assigned by over half of the classifiers (here, outlier detectors). Furthermore, an ensemble of methods can overcome the uncertainty that comes from the unlabeled data. There are two groups of methods that can be used in this layer: (1) ensemble outlier detection algorithms and (2) classification methods that can be used in order to combine the labels of the outlier detectors.

We selected LODA [31] and LSCP [32] as the ensemble outlier detection methods and implemented a major voting as well as a hypergraph-based kernel method as the label combiners. The major voting algorithm receives the labels provided by five outlier detection algorithms (i.e., KNN, HBOS, IOD, CLBOF, and LOF). It assigns the "outlier" label to the samples if the total number of outlier labels assigned by the algorithms exceeds 3.

In addition to majority voting, we employed a hypergraph-based kernel method in order to combine the labels. Kaveh-Yazdy *et al*. [33] have reported that the hypergraph-based kernel successfully models the behavior proximity in the movie recommender systems and predicts more accurate scores. Similarly, we defined a hypergraph in which the vertices were the samples, and the hyperedges were pairs of $\langle A, L \rangle$, where A denotes an outlier detection algorithm, and $L \in \{0,1\}$ is its assigned label.

Subsequently, in an ensemble of n outlier detectors, $2 \times n$ hyperedges could be defined.

The hypergraph adjacency matrix can be filled as:

$$H\left(v_i, \left\langle A_j, L \right\rangle\right) = H\left(v_i, e_{jl}\right) = \begin{cases} 1 & if\ v_i \in e_{jl} \\ 0 & o.w. \end{cases} \quad (4)$$

Where v_i is the i^{th} vertex and e_{jl} denotes the j^{th} hyperedge of label L. This means that the i^{th} sample, v_i, gets label L, which is generated by the algorithm A_j. Let D_v denote the diagonal matrix that includes the total number of hyperedges of vertices (a.k.a. vertex degree matrix), as follows:

$$D_v = \{d_{ii}\}_{i \in \{1,2,\ldots,|V|\}} = \left\{ \sum_{k=1}^{|\varepsilon|} H\left(v_i, e_k\right) \right\} \quad (5)$$

Let L denote the Laplacian of the hypergraph as:

$$L = D_v - HH^T \quad (6)$$

where H and D_v indicate the hypergraph adjacency and vertex degree matrices, respectively. The normalized form of the Laplacian matrix that is used in kernel computing can be calculated as follows:

$$\tilde{L} = D_v^{-\frac{1}{2}} L D_v^{-\frac{1}{2}} \quad (7)$$

The kernel of the hypergraph, denoted by K, is the power series of the normalized Laplacians, i.e. $\tilde{L}$, weighted by the eigenvalues of them, i.e.:

$$K = \sum_{i=1}^{m} \lambda_i (\tilde{L})^i \tag{8}$$

Yu *et al*. [34] have generalized the normalized Euclidean distance between the objects to its equivalent term in the Hilbertian space using hypergraph kernel, as:

$$d\left(\phi(x), \phi(y)\right) = K(x,x) - 2K(x,y) + K(y,y) \tag{9}$$

The computed kernel can be embedded in a kernel-based classification method. In this work, we embedded the computed kernel in a single class support vector machine (SVM) that assigned the outlier labels. Finally, the results of the ensemble layer were evaluated using the provided metrics.

5. Experimental Results and Discussion

Due to the remarked data incompleteness, we decided to solve the problem in a semi-supervised manner. Table 1 shows the total number of records labeled as either normal or outlier in both record sets. The results obtained, shown in Table 1, depict that 28-69% of the suspicious meters are labeled as the normal meters. It seems that more than 1/3 of the water meters are replaced without an anomalous history. Besides the freezing/breaking cases, the unknown decision criteria may justify a large number of meters replaced without a meaningful observation of the anomalous behaviors. As mentioned earlier, the meter failure detection process relies upon the reports by the meter readers, meter consumers, and billing system alerts. The billing system warns the customer desk operator in case of registering the consumption values less than a pre-defined proportion of the yearly average consumption. Unfortunately, the billing system default policy is not defined based on an in-depth field investigation.

Furthermore, the YWWC's billing system is not aware of the demographic data of consumers and their consumption behaviors, as well as the cultural inclinations that affect the behaviors of the consumers in different time windows of the year. Moreover, the meteorological parameters such as the temperature, humidity, and the number of sand storms have not been taken into account to decide about the unsubstantiated lower consumption cases. We applied the trained models to the clean record sets as well.

The results tabulated in Table 1 show that less than 10% of the clean records present anomalous behaviors, which means that these water meters must be considered for future replacement.

The actual ratio of water meters that are not considered for the replacement cannot be determined without a deeper understanding of the random sampling method used to generate the clean record set. Therefore, we propose that (1) the trained models are applied to the whole records in the billing repository in order to discover the potential replacement cases, and (2) proper decision criteria based on the statistical properties of the anomalous records are designed in order to search the billing records periodically.

It is worth reminding that our dataset does not include the ground truth labels, and to this point, we provided the raw results of labeling that make it impossible to compare different methods. Thus we employed the internal metrics in order to evaluate the performance of algorithms in separating the outlier samples from the normal ones.

Nowak-Brzezińska and Horyń [35] have compared seven internal metrics to compare the unsupervised outlier detection methods that are analogous to clustering. They concluded that the internal metrics such as the Dunn Index (DI) and Davies-Bouldin Index (DBI) were able to determine an appropriate number of clusters and enhance the clustering efficiency by representing the outlier's distribution.

Table 1. Total number of samples labeled as "Outlier" or "Normal" in the clean and suspicious record sets.

#	Model	Clean record set		Suspicious record set	
		Normal	Outlier	Normal	Outlier
1	KNN	201889 (99.19 %)	1643 (0.81 %)	12450 (27.85 %)	32256 (72.15 %)
2	HBOS	196363 (96.48 %)	7169 (3.52 %)	13721 (30.69 %)	30985 (69.31 %)
3	IOD	195855 (96.23 %)	7677 (3.77 %)	15128 (33.84 %)	29578 (66.16 %)
4	CLBOF	195851 (96.23 %)	7681 (3.77 %)	15146 (33.84 %)	29560 (66.12 %)
5	LOF	201803 (99.15 %)	1729 (0.85 %)	12582 (28.14 %)	32124 (71.86 %)
6	LODA	190287 (93.49 %)	13245 (6.51 %)	30901 (69.12 %)	13805 (30.88 %)
7	LSCP	187568 (92.16 %)	15964 (7.84 %)	28144 (62.95 %)	16562 (37.05 %)
8	Major voting	184946 (90.87 %)	18586 (9.13 %)	28428 (63.59 %)	16278 (36.41 %)
9	Hypergraph kernel	196183 (96.38 %)	7349 (3.62 %)	23916 (53.49 %)	20790 (46.51 %)

Table 2. Dunn and Davies-Bouldin indices of outlier detection method in the clean and suspicious record sets.

#	Model	Dunn index		DB index	
		Clean record set	Suspicious record set	Clean record set	Suspicious record set
1	KNN	1.649E-4	0.0031	4.572	1.398
2	HBOS	0.0036	0.0032	3.541	1.274
3	IOD	0.0037	0.0029	3.464	1.418
4	CLBOF	0.0034	0.0030	2.478	1.398
5	LOF	1.637E-4	0.0031	4.569	1.398
6	LODA	**0.0036**	**0.0037**	**1.158**	**1.137**
7	LSCP	1.650E-4	1.613 E-4	2.500	1.345
8	Major voting	1.64E-4	0.0030	3.206	1.341
9	Hypergraph kernel	1.526E-4	0.0021	3.507	1.520

Although they reported five other metrics, we selected DI and DBI that did not require contamination ratio or cluster labels. Table 2 shows the DI and DBI values for the selected algorithms. The higher the DI value is, the more cohesive and well-performed clusters are generated. In the opposite direction, the smaller values of DBI lead to better results. With respect to the results tabulated in Table 2, the LODA ensemble outlier detection method performs better than the other methods, and the hypergraph-based ensemble is the second-best detector. LODA is a lightweight ensemble outlier detection that benefitted from defining equal-width histograms that span dimensions separately. Pevný [31] has noted that the KNN and LOF methods are appropriate outlier detectors for the problems with a few numbers of scattered anomalies. In contrast, Isolation Forest, LODA, and one-class SVM have a higher performance in the problems containing clustered outliers with more than 1% contamination rate.

According to the results tabulated in Table 2, the hypergraph-based detector that uses KNN, HBOS, IOD, CLBOF, and LOF has a higher performance than KNN and LOF, implying that the outliers are clustered. According to the results obtained, the Dunn and DB indices of the hypergraph-based method are lower than IOD and LODA and are close to KNN and LOF, which supports the hypothesis that KNN and LOF affect its performance more than HBOS, IOD, and CLBOF. In order to investigate this hypothesis, we re-evaluated the hypergraph kernel with two different configurations, as follows:

- $\langle HBOS, IOD, CLBOF \rangle$
- $\langle HBOS, IOD, CLBOF, LODA \rangle$

The results of our experiments (Ref.: Table 3 and Table 4) support the hypothesis that the performance of the hypergraph-based kernel depends on its embedded methods. The interesting point is that the performance of the hypergraph-based kernel method with an appropriate set of embedded methods is higher than its best-performed embedded method.

Accordingly, in case of the inappropriate method selection, the hypergraph-based ensemble performs worse than the worst method (in Table 2, hypergraph-based ensemble's Dunn and DB indices are worse than LOF). With respect to the results obtained, the proposed framework is designed to use the hypergraph-based ensemble method with the $\langle HBOS, IOD, CLBOF, LODA \rangle$ configuration.

Table 3. Total number of samples labeled as "Outlier" or "Normal" by the hypergraph-based ensemble method with new configurations.

#	Model	Clean record set		Suspicious record set	
		Normal	Outlier	Normal	Outlier
1	Hypergraph kernel with $\langle HBOS, IOD, CLBOF \rangle$ core	184043 (90.43 %)	19489 (9.57 %)	27806 (62.20 %)	16900 (38.80 %)
2	Hypergraph kernel with $\langle HBOS, IOD, CLBOF, LODA \rangle$ core	188552 (92.64 %)	14980 (7.36 %)	28911 (64.67 %)	15795 (35.33 %)

Table 4. Dunn and Davies-Bouldin indices of the hypergraph-based ensemble method with new configurations.

#	Model	Dunn index		DB Index	
		Clean record set	Clean record set	Clean record set	Suspicious record set
1	Hypergraph kernel with $\langle HBOS, IOD, CLBOF \rangle$ core	0.0031	0.0031	2.184	1.380
2	Hypergraph kernel with $\langle HBOS, IOD, CLBOF, LODA \rangle$ core	**0.0038**	**0.0040**	**1.123**	**1.110**

6. Conclusions

In this work, we aimed to study the problem of water meter replacement in the Yazd municipal water network. So far, the replacement decision has been triggered by the manual reports and the YWWC billing system alerts. Despite the automatic alerting, there has not been any deterministic decision-making process.

The replacement decision is made based on the results of a comparison between the consumption values and a pre-defined threshold computed based on the proportion of yearly average consumption. We designed a smart recommender system that was able to detect anomalous behaviors in the registered records extracted from the YWWC billing system. The detection process could be run by request or periodically and implemented based on an ensemble of the semi-supervised outlier detection methods.

The results of our tests showed that the best-performed proposed framework labeled more than 60% of the replaced meters with the "Normal" value. Furthermore, it discovers a small set of water meters that present an anomalous behavior but do not replace their history. However, the problem of water meter replacement is widely studied with respect to the meters' mechanical properties. Our work is one of the pioneering studies that take into account the consumption history alone. The proposed recommender system is much less costly than the periodic inspections to detect the faulty meters. Besides, it can be extended to extract more complex anomalous patterns. The most similar work to ours is the Casini *et al*. [36] deep-water framework. They used 15M consumption records, and labeled them as valid/in-valid with experts' help, and trained a deep neural network to predict the water meter failure. In contrast to Casini *et al*.'s [36] framework, the proposed framework does not require the labeled data. Furthermore, it uses an explainable approach that facilitates further mechanical investigations. Our experiments demonstrated that applying ensemble outlier detection did not necessarily increase the accuracy of the methods because it may be affected by the methods that were not suitable for the problem. Based on this fact, we suggest that the researchers plan to test the performance of different outlier methods before combing them into an ensemble framework.

As future work, we intend to design a fast test with the minimum computational load to determine the best replacement candidates using Minimum Covariance Determinant Estimator (MCDE) methods [41]. In addition, we recommend YWWC to carry out the post-replacement flow tests for unbroken/unfrozen water meters as well as registering the reason for replacements. The mentioned tests present the statistical properties of the erroneously registered values. We plan to implement a numerical model in order to reconstruct the original consumption values approximately.

References

[1] H. Alegre *et al.*, *Performance Indicators for Water Supply Services-Second Edition*, 2nd Edition, Vol. 12. IWA Publishing, 2013.

[2] A.O. Lambert, "International Report: Water losses management and techniques," *Water Supply*, Vol. 2, No. 4, pp. 1–20, Sep. 2002.

[3] F. Arregui, F. Gavara, J. Soriano, and L. Pastor-Jabaloyes, "Performance Analysis of Ageing Single-Jet Water meters for Measuring Residential Water Consumption," *Water*, vol. 10, No. 5, p. 612, May 2018.

[4] V. Puleo, C.M. Fontanazza, V. Notaro, M. De Marchis, G. La Loggia, and G. Freni, "Definition of Water meter Substitution Plans based on a Composite Indicator," *Procedia Eng.*, Vol. 70, pp. 1369–1377, Jan. 2014.

[5] L.E. Orr, V.A. Enna, and M.C. Miller, "Analysis of a Water-meter Replacement Program," *J. Am. Water Work. Assoc.*, Vol. 69, No. 2, pp. 68–71, Feb. 1977.

[6] D.M. Stoker, S.L. Barfuss, and M.C. Johnson, "Flow measurement accuracies of in-service residential water meters," *J. Am. Water Works Assoc.*, Vol. 104, No. 12, pp. E637–E642, Dec. 2012.

[7] M.R. Clinciu and R. Clinciu, "Determining the predictive reliability of the water meters using the failure rate estimation. Case study-single jet cold water meter," *IOP Conf. Ser. Mater. Sci. Eng.*, Vol. 399, p. 012008, Sep. 2018.

[8] F.J. Arregui, E.J. Cabrera, R. Cobacho, and J. García-Serra, "Key factors affecting water meter accuracy," in *Proceedings of the IWA Leakage Conference*, Halifax, 2005, pp. 1–10.

[9] F.J. Arregui, M. Balaguer, J. Soriano, and J. García-Serra, "Quantifying measuring errors of new residential water meters considering different customer consumption patterns," *Urban Water J.*, Vol. 13, No. 5, pp. 463–475, Jul. 2016.

[10] C.M. Fontanazza, G. Freni, G. La Loggia, V. Notaro, and V. Puleo, "A composite indicator for water meter replacement in an urban distribution network," *Urban Water J.*, Vol. 9, No. 6, pp. 419–428, Dec. 2012.

[11] F. Yazdandoost and A. Izadi, “An asset management approach to optimize water meter replacement,” *Environ. Model. Softw.*, Vol. 104, pp. 270–281, Jun. 2018.

[12] M. A. Cardoso, M.S. Silva, S.T. Coelho, M.C. Almeida, and D.I.C. Covas, “Urban water infrastructure asset management–a structured approach in four water utilities,” *Water Sci. Technol.*, Vol. 66, No. 12, pp. 2702–2711, Dec. 2012.

[13] C.M. Fontanazza, V. Notaro, V. Puleo, and G. Freni, “The apparent losses due to metering errors: a proactive approach to predict losses and schedule maintenance,” *Urban Water J.*, Vol. 12, No. 3, pp. 229–239, Apr. 2015.

[14] R.P. Weller, S.L. Barfuss, and M.C. Johnson, “The effect of surge flows on residential water meters,” *AWWA Water Sci.*, Vol. 1, No. 1, p. e1117, Jan. 2019.

[15] M.N. Kanyama, C. Nyirenda, and N. Clement-Temaneh, “Anomaly Detection in Smart Water metering Networks,” in *The 5th International Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII2017)*, Beijing, 2017, pp. 1–10.

[16] T. Andrysiak, Ł. Saganowski, and P. Kiedrowski, “Anomaly Detection in Smart Metering Infrastructure with the use of Time Series Analysis,” *J. Sensors*, Vol. 2017, pp. 1–15, 2017.

[17] A. Moahloli, A. Marnewick, and J. Pretorius, “Domestic water meter optimal replacement period to minimize water revenue loss,” *Water SA*, Vol. 45, No. 2 April, Apr. 2019.

[18] E. Vonk, D.G. Cirkel, and M. Blokker, “Estimating peak daily water demand under different climate change and vacation scenarios,” *Water (Switzerland)*, Vol. 11, No. 9, p. 1874, Sep. 2019.

[19] M. Ebrahim Banihabib and P. Mousavi-Mirkalaei, “Extended linear and non-linear auto-regressive models for forecasting the urban water consumption of a fast-growing city in an arid region,” *Sustain. Cities Soc.*, Vol. 48, p. 101585, Jul. 2019.

[20] N. Rasifaghihi, S.S. Li, and F. Haghighat, “Forecast of urban water consumption under the impact of climate change,” *Sustain. Cities Soc.*, Vol. 52, p. 101848, Jan. 2020.

[21] G.-F. Fan, S. Qing, H. Wang, W.-C. Hong, and H.-J. Li, “Support Vector Regression Model based on Empirical Mode Decomposition and Auto Regression for Electric Load Forecasting,” *Energies*, Vol. 6, Nno. 4, pp. 1887–1901, Apr. 2013.

[22] Y. Chen, W.-C. Hong, W. Shen, and N. Huang, “Electric Load Forecasting based on a Least Squares Support Vector Machine with Fuzzy Time Series and Global Harmony Search Algorithm,” *Energies*, Vol. 9, No. 2, p. 70, Jan. 2016.

[23] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, *Morgan Kaufmann Publishers Inc.*, 3rd Ed., San Francisco, CA, 2011.

[24] Y. Li, M. Dong, and Y. Ma, “Feature selection for clustering with constraints using Jensen-Shannon divergence,” *in Proceedings of the 19th International Conference on Pattern Recognition*, Tampa, Fl. Dec. 2008, pp. 1–4.

[25] R. Guzmán-Martínez and R. Alaiz-Rodríguez, “Feature Selection Stability Assessment based on the Jensen-Shannon Divergence,” *Springer*, Berlin, Heidelberg, 2011, pp. 597–612.

[26] F. Bagattini, I. Karlsson, J. Rebane, and P. Papapetrou, “A classification framework for exploiting sparse multi-variate temporal features with application to adverse drug event detection in medical records,” *BMC Med. Inform. Decis. Mak.*, Vol. 19, No. 1, p. 7, 2019.

[27] J. Lin, E. Keogh, L. Wei, and S. Lonardi, “Experiencing SAX: a Novel Symbolic Representation of Time Series,” *Data Min. Knowl. Discov.*, Vol. 15, No. 2, pp. 107–144, 2007.

[28] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, “A symbolic representation of time series, with implications for streaming algorithms,” *in Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, San Diego, CA, 2003, p. 2-11.

[29] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, “Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median,” *J. Exp. Soc. Psychol.*, Vol. 49, No. 4, pp. 764–766, Jul. 2013.

[30] M. Goldstein and S. Uchida, “A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data,” *PLoS One*, Vol. 11, No. 4, p. e0152173, Apr. 2016.

[31] T. Pevný, “Loda: Lightweight on-line detector of anomalies,” *Mach. Learn.*, vol. 102, No. 2, pp. 275–304, Feb. 2016.

[32] Y. Zhao, Z. Nasrullah, M.K. Hryniewicki, and Z. Li, “LSCP: Locally Selective Combination in Parallel Outlier Ensembles,” in *Proceedings of the 2019 SIAM International Conference on Data Mining*, Philadelphia, PA, 2019, pp. 585–593.

[33] F. Kaveh-Yazdy, X. Kong, J. Li, F. Li, and F. Xia, “Customer Rating Prediction using Hypergraph Kernel based Classification,” in *The 2013 International Conference on Active Media Technology (AMT’13)*, Maebashi, Japan, 2013, pp. 187–192.

[34] K. Yu, L. Ji, and X. Zhang, “Kernel Nearest-Neighbor Algorithm,” *Neural Process. Lett.*, Vol. 15, pp. 147–156, 2002.

[35] A. Nowak-Brzezińska and C. Horyń, “Outliers in rules-the comparison of LOF, COF, and KMEANS algorithms.,” *Procedia Comput. Sci.*, Vol. 176, pp. 1420–1429, 2020.

[36] L. Casini, G. Delnevo, M. Roccetti, N. Zagni, and G. Cappiello, “Deep Water: Predicting Water Meter Failures Through a Human-Machine Intelligence Collaboration,” *in Proceedings of the Human Interaction and Emerging Technologies*, T. Ahram, R. Taiar, S. Colson, and A. Choplin, Eds. Cham, Switzerland: Springer International Publishing, 2020, pp. 688–694.

[37] A. Eftekhar, W. Juffali, J. El-Imad, T.G. Constandinou, and C. Toumazou, “Ngram-Derived Pattern Recognition for the Detection and Prediction of Epileptic Seizures,” *PLoS One*, Vol. 9, No. 6, p. e96235, Jun. 2014.

[38] R. Hernández-García, J. Ramos-Cózar, N. Guil, E. García-Reyes, and H. Sahli, “Improving Bag-of-Visual-Words model using visual n-grams for human action classification,” *Expert Syst. Appl.*, Vol. 92, pp. 182–191, Feb. 2018.

[39] J. Violos, K. Tserpes, I. Varlamis, and T. Varvarigou, “Text Classification using the N-Gram Graph Representation Model over High Frequency Data Streams,” *Front. Appl. Math. Stat.*, Vol. 4, p. 41, Sep. 2018.

[40] E. Kermany, H. Mazzawi, D. Baras, Y. Naveh, and H. Michaelis, “Analysis of Advanced Meter Infrastructure Data of Water Consumption in Apartment Buildings,” *in Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Chicago, IL, 2013, pp. 1159–1167.

[41] M. Mohammadi and M. Sarmad, “Outlier Detection for Support Vector Machine using Minimum Covariance Determinant Estimator,” *J. AI Data Min.*, Vol. 7, No. 2, pp. 299–309, 2019.

Appendix A

In this section, we present an N-gram-based visualized investigation that targets to illustrate the differences between the sequences of consumption values in the clean record set versus the suspicious record set.

The N-gram models traditionally used in the textual data analysis problems are able to discover the frequent sequential patterns [37]–[39]. N-grams, as subsequences of length N, are used in order to represent the subsequences of the streams and text documents. Suppose that a document is repeated several times in a text repository. Then the subsequences of the document are repeated more than the subsequences of the other documents. Consequently, studying the frequent N-grams in a repository can reveal the prevalent sequential patterns. Moreover, the distinction between two sequential repositories can be inferenced by taking into account the diversity of their N-grams.

We extract the N-grams of the symbolic consumption time series, and demonstrate a limited number of the most frequent N-grams (e.g. 200 N-grams), which are worthwhile to present the distinction between the clean and the suspicious record sets in Figure 3.

The right subplot in this figure shows the distribution of symbols in different positions of the most frequent N-grams extracted from the clean record set. The left subplot shows the same chart generated using the suspicious records. The sunburst plots in Figure 3 are made up of three layers in which the *i*th layer shows the distribution of the symbols in the *i*th position of the N-grams.

Figure 3. Sunburst plot of the 200 most frequent N-grams in clean record set (right) and suspicious record set (left).

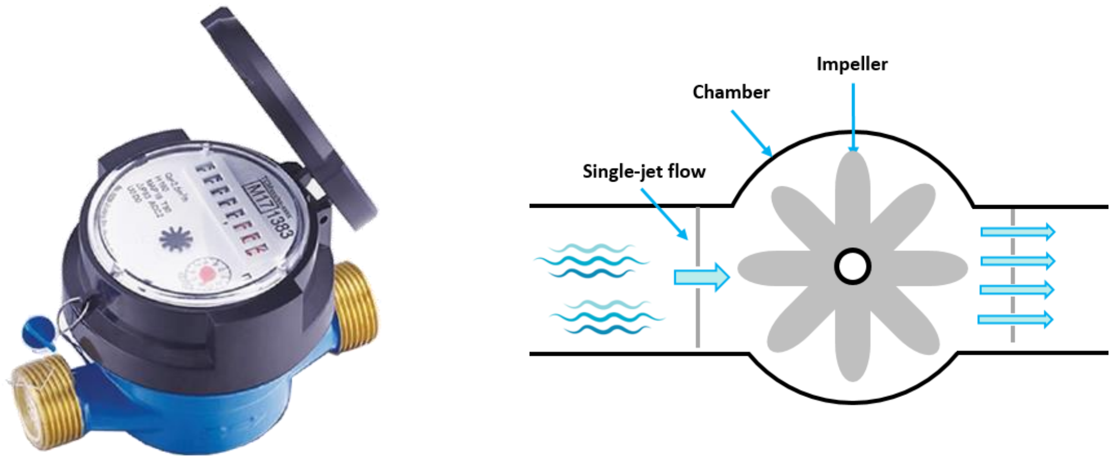

Figure 4. A single-jet water meter (left) and a schematics view of its structure (right).

The areas of the sectors in the layers are proportional to the conditional probability of the corresponding symbol in its associated position. The mentioned sunburst plots show how the most frequent N-grams in each record set are formed.

The differences between the subplots present the differences between the distributions of the symbols in their positions. The most attractive difference is repeating the same symbol in all three positions. In addition, the second and third most-repeated symbols after every symbol in the first layer of two subplots are different, which shows that distinct trends of values in time series exist. The existing list of the same symbols means that the meters that are mostly of single-jet type are corrupted and lose their sensitivity. In such meters, a single cavity lets the water flows inside the chamber and rotates the impeller, as shown in Figure 4. When the meter's sensitivity is reduced, the small and surge flows are not measured precisely; hence, the registered consumption values bias to the values in a limited range. Similarly, Kermany *et al*. [40] have reported that they have labeled the zero-consumption meters as the faulty meters and have removed their records from their data.

ارائه توصیه‌ی تعویض کنتور آب برای شبکه‌های توزیع آب شهری با استفاده از روش‌های جمعی تشخیص موارد پرت

فاطمه کاوه یزدی و سجاد ظریفزاده*

دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران.

ارسال ۲۰۲۱/۰۴/۰۱؛ بازنگری ۲۰۲۱/۰۵/۰۷؛ پذیرش ۲۰۲۱/۰۶/۲۶

چکیده:

کنتورهای آب بنا به ساختار و شرایط استفاده همواره در معرض فرسایش، شکستگی، یخزدگی و نشت هستند. در این راستا، مطالعات متعددی با هدف تعیین زمان مناسب برای تعویض کنتورهای فرسوده انجام شده‌اند. مطالعات متقدم، از ویژگی‌های متنوعی از جمله شرایط آب‌وهوایی، شرایط استفاده، فشار شبکه توزیع و ساختار کنتورها برای پیش‌بینی فرسودگی استفاده می‌نمایند. این مقاله به معرفی یک چهارچوب توصیه‌گر می‌پردازد که با دریافت مقادیر مصرف آب، توصیه لازم مبنی بر تعویض کنتور فرسوده را ارائه می‌نماید. این چهارچوب، با پیش‌پردازش سری‌های زمانی مقادیر مصرف آب ثبت شده، ویژگی‌های موثر را استخراج می‌کند و سپس با استفاده از چند روش جمعی بدون نظارت/نیمه نظارتی که برای کشف موارد پرت مورد استفاده قرار می‌گیرند برچسب‌های طبیعی و پرت را به کنتورها اعطا می‌کند. در آخرین مرحله، یک روش جمعی مبتنی بر رای‌گیری ابرگراف تمام برچسب‌های اختصاص داده شده توسط روش‌های مختلف را دریافت نموده و با ترکیب آنها به یک برچسب نهایی می‌رسد. به دلیل عدم دسترسی به دادگان دارای برچسب حقیقی، برای ارزیابی روش پیشنهادی و روش‌های مورد مقایسه از نرخ مثبت کاذب و نمایه‌های Dunn و Davies-Bouldin بهره‌گرفته شده است. نتایج آزمایشات مقایسه‌ای نشان داده‌اند، چهارچوب پیشنهادی می‌تواند خوشه‌های فشرده با واریانس کمتر تولید کند.

کلمات کلیدی: اندازه‌گیری آب، تلفات ظاهری، تشخیص شکست، تشخیص موارد پرت، چندجزئی، تحلیل سری‌های زمانی.