

High-Resolution Tap-Based IoT System for Flow Data Collection and Water End-Use Analysis

Man-Ho Luk, Cheuk-Wang Yau, Philip W. T. Pong, *Senior Member, IEEE*, Angela P. Y. Lee, Edith C. H. Ngai, *Senior Member, IEEE*, and King-Shan Lui

Abstract—Knowledge on water-use patterns in residential settings can help policymakers formulate well-targeted water conservation measures and evaluate the efficacy of such measures. Prior studies have applied machine learning techniques to disaggregate household-level water consumption data, collected by smart meters, into specific end-uses categories, such as showering, basin use, kitchen use and use by washing machine. However, the analysis of the effect of sampling interval on classification accuracy level remains an under-investigated issue. This paper seeks to fill this knowledge gap by identifying an optimal sampling interval that can achieve a high level of classification accuracy while overcoming constraints of on-device data storage, data transmission, and energy consumption. To understand the benefits of collecting fine-grained data for machine learning, we have built a High-Resolution Tap-Based IoT System comprising a set of Wi-Fi-based Tap Sensors, gateway infrastructure and a secure data processing pipeline. Based on empirical tap-based data collected over an eight-month period, we concluded that when the sampling interval decreases slightly from 5 seconds to 1 second, the accuracy level of the end-use classification model increases significantly from 66.6 % to 76.1 %. This paper also highlights the challenges of deploying IoT sensors to collect water consumption data in a domestic setting. In order to collect sufficient ground truth data for the training and verification of a generalizable water end-use disaggregation model, it is necessary to sophisticate the flow data collection system by adopting low-power wide-area network technologies and reducing the level of energy consumption of the flow sensing components.

Index Terms—Internet of Things, sensor systems, flowmeters, smart meters, embedded systems, machine learning.

I. INTRODUCTION

RESEARCH found that while cities around the world have experienced a gradual decline in per capita domestic water consumption level in the last two decades, Hong Kong has been going against the international trend [1]. To tame the growth in the level of domestic water consumption, an effective water conservation policy, which is based on a nuanced understanding of how much water and for what purposes people are using, is needed. Such information, however, have never been collected systematically in Hong

Kong. *Water auditing*, a conventional method to inventory all water usage in a single facility and identify ways to increase water-use efficiency, could be applied at the urban scale to help determine the amount of unaccounted-for-water and non-revenue water in a city-wide water distribution system [2]. We can also apply water auditing at the household level to help understand the composition of domestic water consumption (i.e., the amount of water-use in each *end-use category*, such as kitchen use, basin use, showering, and use by washing machine) [3].

One possible way to collect data on the composition of water-use in a domestic setting is to install a flow sensor (also known as a flowmeter), together with a data logger, at each water outlet. By doing so, the questions of “how much water have been used” and “where did one use the most water” could be addressed. While this approach may work for a small number of households, it is not a cost-effective option for large-scale roll-out. Fortunately, emerging machine learning technologies, as demonstrated in studies such as [4], have enabled us to label the end-use category of each water-use event recorded by the utility’s smart water meter efficiently. Very often, water end-use disaggregation is divided into two sub-problems: *Event disaggregation* and *end-use classification* [5]. Overlapping water-use events collected by smart water meters that monitor household-level water use are first disaggregated into multiple single-use events. These single-use events are then classified into different water end-use categories. Since concurrent water-use from multiple water taps is infrequent in domestic settings, end-use classification plays a large part in understanding of the composition of domestic water consumption.

Results from previous studies showed that the training of machine learning algorithms is feasible by using simulated water consumption data or secondary data collected from prior research on patterns of domestic water end-use. For instance, Wonders *et al.* used synthesized data [6]; most of the recently developed algorithms [7]–[9] relied on the dataset generated by the South East Queensland residential end-use study in 2011 [10]. However, the collection of primary data or ground truth data can guarantee an accurate labeling of the end-use category of all water-use events, and thus increase the level of confidence when developing and validating more advanced water end-use disaggregation techniques. This paper describes our efforts in building and deploying our own Internet of Things (IoT) sensors system to pave the way for conducting *Smart Water Auditing* for Hong Kong.

Smart Water Auditing is a novel IoT application specialized

This work was supported in part by Philomathia Foundation and WYNG Foundation.

Man-Ho Luk, Edith C. H. Ngai (*corresponding author), and King-Shan Lui are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: luk10099@connect.hku.hk; chngai@eee.hku.hk; kslui@eee.hku.hk).

Cheuk-Wang Yau and Angela P. Y. Lee are with the Centre for Water Technology and Policy, The University of Hong Kong, Hong Kong (e-mail: mcwyau@hku.hk; angela01@hku.hk).

Philip W. T. Pong is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: philip.pong@njit.edu).

in collecting and analyzing residential water usage data. The water sector is starting to adopt IoT technologies to help manage the water distribution system and wastewater treatment [11], and to facilitate metering and demand management [12], [13]. However, collection of empirical data on tap-based water usage, using IoT sensor systems, in participants' homes requires additional consideration. Harnessing cloud and machine learning technologies, empirical data collected are then used to train machine learning algorithms for water end-use disaggregation. Our ultimate goal is to utilize labeled, tap-based data collected from a small number of households to train a generalizable machine learning model for classifying and labeling the end-use category of each unknown water-use event collected from any household accurately. We envision that, with a wider coverage of capable smart water meters and with an accurate end-use disaggregation model available, water utilities can conduct *Smart Water Auditing* to learn about the water usage patterns of their customers and use such data for evidence-based water demand management and forecasting. A continuous stream of real-time water-use data can also be used for developing and sending personalized messages to water users to motivate water conservation. Fig. 1 presents a high-level overview of what *Smart Water Auditing* is, with reference to the five-layer IoT architecture [14], and outlines the technology and protocol stacks involved. To the best of our knowledge, our proposed IoT sensors system is the first domestic water-use data collection system specifically designed for developing and enhancing machine learning algorithms for water end-use disaggregation.

The rest of the paper is organized as follows: Section II reviews the relevant literature on domestic water end-use disaggregation, and summarizes the limitations of previous studies. Section III outlines the data collection requirements, as well as the system design of our IoT sensors system for collecting ground truth water consumption data for training of Machine Learning models. Section IV discusses the challenges and lessons of deploying the data collection system in a domestic setting. We have used the empirical data collected to investigate the effect of sampling interval on the level of classification accuracy of water end-use disaggregation models. Preliminary findings are presented in Section V. Section VI concludes our work and recommends directions for future work in this area.

II. RELATED WORK

Cominola *et al.* [15] and Roshan and Kumar [16] have reviewed domestic water end-use studies conducted since the 1990s. The emergence of digital flowmeter and Automatic Meter Reading (AMR) systems has enabled water utilities and researchers to acquire time series data on household water consumption for end-use analysis. Table I compares the data collection method and technical specifications of this study with other studies on domestic water end-use analysis.

Earlier studies made use of software packages, such as *Trace Wizard*, *Identiflow*, and *BuntBrainForEndUses*, which are commercial software packages that have incorporated a decision tree algorithm for conducting water end-use analysis [5], [15].

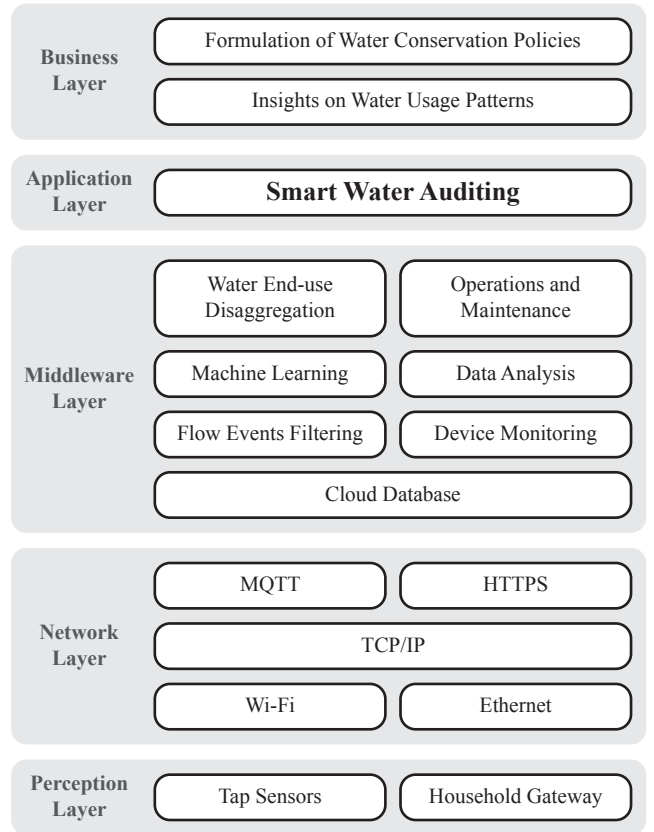


Fig. 1. High-level overview of technologies and protocols used in *Smart Water Auditing*.

Trace Wizard has achieved an average classification accuracy level of 70 % [23]; however, concurrent usage contributed by simultaneous water-use at multiple taps would reduce its accuracy [8]. *Identiflow* claimed a higher accuracy level of 74.8 % [23]. *Identiflow* extracts the physical characteristics of each water-use device (e.g., volume, flow rate, duration, etc.) for classification. Water-use events sharing similar physical characteristics will be classified into the same end-use category. However, the accuracy of *Identiflow* is affected in case a household is equipped with two or more devices that share very similar physical characteristics. [15] In addition, using these decision tree-based software packages is labor-intensive and time-consuming, as they require gathering of information on the characteristics of different fixtures. Occasional manual annotation of uncommon or unusual water-use events is also required. Hence, it is cost-prohibitive to scale up and support data analysis using these tools in studies with a large sample size [5].

Recent studies have focused on applying machine learning to facilitate fully automatic and more scalable water end-use disaggregation [4], [6]–[8], [24]. The development of a machine learning model, in general, consists of two phases: 1) The collection of training data; and 2) The extraction of features from training data for training of the machine. We first review what training data were collected, and how they were collected, in existing work.

TABLE I
COMPARISON OF THIS WORK WITH PREVIOUS DOMESTIC WATER END-USE STUDIES

Reference ^a	Year	Location	Metering Granularity	Volumetric Resolution ^b (mL)	Sampling Interval (s)	Period of Data Collection	Sample Size (households)
This work	2020	Hong Kong	Tap-based	2	1	8 months	3
[5]	2018	<i>Unspecified</i>	<i>Unspecified</i>	60	3	2 weeks	10
[17]	2016	USA	Household-level	47	10	2 weeks	762
[4]	2016	USA	Tap-based	<i>Unspecified</i>	60	7 months	8
[18]	2016	Australia	Household-level	500	60	2 to 5 weeks	68
[19]	2011	USA	Household-level	<i>Unspecified</i>	10	1 month	209
[20]	2011	USA	Household-level	38	10	<i>Unspecified</i>	735
[10]	2011	Australia	Household-level	14	5	10 weeks	200+
[21]	2011	Australia	Household-level	<i>Unspecified</i>	5	2 months	93
[22]	2009	Australia	Household-level	14	10	<i>Unspecified</i>	200

^a The studies are sorted by publication date in descending order.

^b Converted from number of pulses per unit volume (rounded up to the nearest mL).

Most of the previous studies have collected domestic water consumption data using commercial flowmeters. These flowmeters are usually in-line meters that record the total amount of water-use at the household level. Whenever water flows through the flow meter, the meter generates a pulse which is then captured by a data logger; the data logger is typically capable of recording the accumulated volume with sampling intervals less than one minute. These commercial flowmeters have very high precision level (as high as 14 milliliters (mL) per pulse [22]). However, since they are primarily designed for billing purposes, most of them cannot support sampling interval shorter than 5 s. A lack of finer-grained training data poses a major challenge to developing and improving machine learning algorithms for water end-use disaggregation. This is because fine-grained details, such as changes in flow rate, are lost or distorted when a dataset with longer sampling interval is used for reconstructing the temporal pattern of a water-use event (i.e., a flow profile).

Fine temporal data granularity is critical for an accurate representation of how water is used in a domestic setting. Consider the example of filling a 250 mL cup. The cup is filled up in 2.5 s using at a bathroom basin tap with a typical flow rate of 6 L/min (100 mL/s). For instance, when the sampling interval is set at 5 s, the average flow rate of this 2.5 s cup-filling event is reduced by 50 % to 50 mL/s. This distorted flow profile could lead to erroneous event disaggregation and false end-use classification by the machine learning algorithms. Analysis of the effect of sampling interval on the level of classification accuracy remains an under-investigated issue. This paper seeks to fill the knowledge gap by identifying an optimal sampling interval that can maintain a high level of classification accuracy while overcoming the constraints of on-device data storage, communication, and energy budgets.

The use of commercial flowmeters to measure tap-based water consumption data is not feasible as they are too bulky

and costly to be installed at each water outlet of each home. A majority of existing work have used decision tree-based tools like *Trace Wizard* to break down household-level data into individual water-use events and classify them using salient features such as volume, duration, and flow rate [10], [17], [19]–[21]. Despite the wide application of *Trace Wizard* to analyze flow trace data, the accuracy level of *Trace Wizard* could not be ascertained in the absence of ground truth data. Evaluation of the reliability of machine learning predictive models against detailed water consumption diaries, like the model verification work conducted by [7], is not perfect, as the accuracy of data is dependant upon the vagaries of study participants' self-reporting skills. Training and validating water end-use disaggregation algorithms using datasets with potentially-misabeled water usage may impact a classifier's actual performance. [4] is one of the few studies that had used empirical tap-based water usage data for verification of machine learning models. The sampling interval set at 60 s, nevertheless, is unlikely to be sufficiently short for providing fine-grained details for machine learning, and the authors did not provide any details regarding the data collection devices used. Since there is no market available solutions for the collection of tap-based water consumption data with a sampling interval of 5 s or shorter, we, thus, have turned to develop our own system to collect the much needed ground truth data.

Next, we review how domestic water consumption data were used to develop machine learning algorithms for water end-use disaggregation in the literature. To help inform the data collection requirements of our proposed system, we focus on what features researchers found most useful for model training. Table II presents the salient features of a water-use event that have been used for model training by other research teams. Total volume of a water-use event is used in all the selected studies; duration and flow rate of the event are used in most of

TABLE II
FEATURES USED FOR WATER END-USE DISAGGREGATION ALGORITHMS
IN PREVIOUS STUDIES

Reference ^a	Water Usage Event Features			Other Features
	Volume	Duration	Flow Rate	
[5]	✓		Average	
[24]	✓	✓		Electricity usage
[8]	✓	✓	Maximum	
[4]	✓	✓		Hour of the day Hot water usage Toilet usage Lights usage
[6]	✓		Average S.D. ^b Mode	
[7]	✓	✓	Maximum Mode F.M. ^c Gradient ^d	

^a The studies are sorted by publication date in descending order.

^b Standard deviation.

^c Frequency of mode.

^d Gradients of initial and final flow rates of an event, and their peak magnitudes.

the studies. These three features describe the most prominent characteristics of a water-use event. Nguyen *et al.* [7] also utilized the gradients of the initial and final flow rates of an event. Vitter and Webber [24] and Gao *et al.* [4] have used electricity consumption data to enhance their classification models: They claimed that electricity usage provides a temporal cue for certain types of water end-use activities (e.g., dish washing and washing clothes). Their models, moreover, have produced decent classification results using water consumption data alone. According to the literature reviewed, we determined that the extraction of the following features from the dataset on domestic water consumption would benefit the development of the machine learning algorithms: 1) Total volume; 2) Duration; 3) Average flow rate; and 4) Gradients of the initial and final flow rates of a water-use event. Simultaneous electricity consumption data can improve classification accuracy, but it is not absolutely necessary.

III. SYSTEM REQUIREMENTS AND DESIGN

The goal of our study is to collect data to develop a machine learning algorithm that breaks down household-level usage data into individual water-use events and to label those events accurately. Prior studies on water end-use disaggregation shed light on the most important features we have to extract from a dataset. To extract these features, the data collection device must have a high level of precision and offer high temporal resolution. As the devices would be installed in bathrooms and kitchens, they must be able to withstand high temperature and high humidity level. Other situational constraints, such as difficulties of providing wired network and mains electricity in bathrooms or under a kitchen sink, pose additional challenges. In this section, we first define the variables to be measured. We

then describe the requirements of our data collection system, and explain our design in detail.

A. Flow Event and Flow Profile

A *flow event* is defined as a discrete water-use incident that takes place at a specific outlet. Let the outlet of where a flow event e_i occurs be o_i . The start time of e_i , denoted as t_i^s , is the moment that water starts flowing out from o_i . The end time of e_i , denoted as t_i^e , is the final moment that water runs from o_i right before the water flow stops. It is important to distinguish a flow event from a *water end-use activity* (e.g., handwashing, showering, etc.). An end-use activity may consist of one or multiple flow events. For instance, if a person turns on and off the tap three times during a shower, three flow events would be registered for this showering event.

A *flow profile*, which is a plot of flow rate of the flow event against time, is used to describe how much water is used at any instant of a flow event. We denote the profile of event e_i as p_i . p_i describes the flow rate in time interval $[t_i^s, t_i^e]$. In practice, a digital data collection device captures p_i as a discrete-time signal based on a finite sampling interval T_s . In other words, the device records the accumulated volume of water-use during the interval every T_s . Fig. 2(a) shows an example of a flow profile with $T_s = 1$ s. The average flow rate at time step t is $p_i[t]$; the total volume of e_i is $\sum_{t=t_i^s}^{t_i^e} p_i[t]$. Mathematically, flow event e_i is defined as (o_i, t_i^s, t_i^e, p_i) .

Let a flow event captured by the household's water meter be e_m . p_m is the superposition of all flow events that take place at one or more outlets during the time period $[t_m^s, t_m^e]$. The problem of event disaggregation is to identify the set of flow events $E_m = \{e_k | t_k^s \geq t_m^s, t_k^e \leq t_m^e\}$ such that $\sum_{e_k \in E_m} p_k[t] = p_m[t]$ where $t \in [t_m^s, t_m^e]$. All outlets in the household form a set Ω . Each of the pre-defined end-use categories is a mutually exclusive subset of Ω . End-use classification is the task of assigning an end-use category C_j to e_k such that $o_k \in C_j$.

B. System Requirements

1) *Sampling Interval*: In Section II, we identified that most market-available data collection devices could only achieve a minimum sampling interval (or a maximum temporal resolution) of around 5 s. We argue that an absence of finer-grained data is a potential obstacle to further improvement of the accuracy of the water end-use classification model.

The sampling interval T_s , set at different temporal resolution, has profound effects on flow profile reconstruction. Fig. 2 presents the flow profiles of the same flow event developed with different T_s . We captured this event from the basin tap of an actual household with $T_s = 1$ s, and down-sampled it to $T_s = 5$ s, 10 s, and 30 s. The flow event starts at $t = 11$ s and ends at $t = 36$ s. When set at the highest temporal resolution (i.e., $T_s = 1$ s), Fig. 2(a) shows that the flow rate reaches a peak for a few seconds after the flow event was started, then drops for around two seconds, and rises, and finally plateaus until the flow event was ended. This flow profile is a graphical representation of the flow profile of a typical basin-use event: A person first turned on the basin tap but found it was running too fast than expected. The person then adjusted the tap to slow

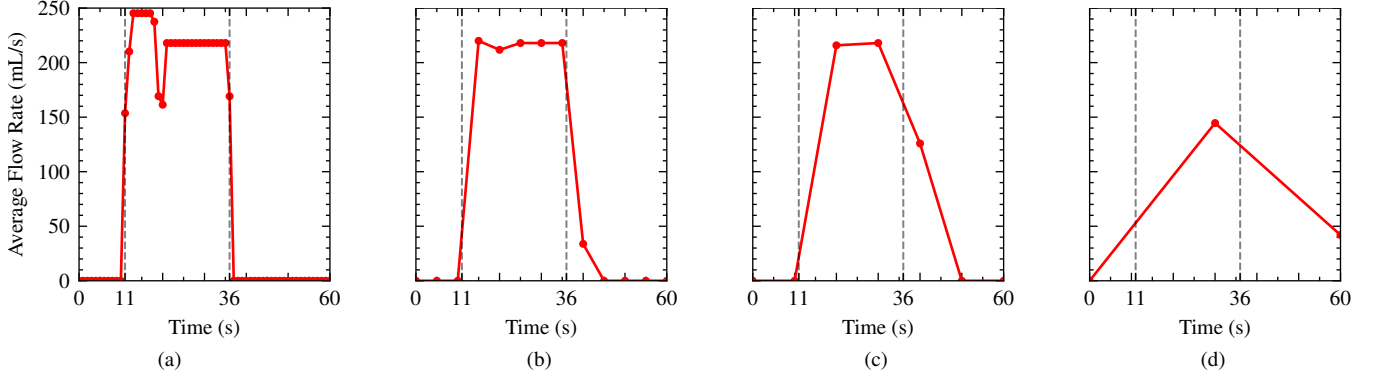


Fig. 2. (a) A sample flow profile captured at a basin tap with a sampling interval of $T_s = 1$ s. The flow profile is down-sampled to (b) $T_s = 5$ s; (c) $T_s = 10$ s; and (d) $T_s = 30$ s to demonstrate the effect of increasing T_s on the shape and magnitude of the flow profile.

down the water flow. Fig. 2(b) to (d) illustrate four different flow profiles with increasing T_s . In these scenarios, we assume that the flow rate is first sampled at $t = 0$, then at multiples of their corresponding T_s . It can be observed that the level of distortion in the shape of the flow profile increases with T_s . The fluctuation of flow rate at around $t = 15$ s is averaged out and is no longer recognizable from the flow profiles with $T_s = 5$ s or above. In addition, the start time and the end time of the flow event are delayed with longer T_s . These examples demonstrated two issues pertaining to inadequate temporal resolution of training datasets. First, the flow profile would be distorted, resulting in a loss of distinctive features of flow rate changes. Second, the failure of capturing the exact start time of the flow event would also result in the distortion of the geometry of the flow profile. Hence, feature extraction by machine learning models would be affected.

Some features (e.g., gradients of the initial and final flow rates of an event) are critical for the development of a machine learning model for water end-use disaggregation [7]. Such features, however, could only be extracted if there is an accurate reconstruction of the flow profile. Hence, we decided to build our own high-resolution data collection system. Instead of counting pulses at a given T_s , we keep track of the pulse duration (time difference between two successive rising edges) of every flowmeter-generated pulse. In this way, we can reconstruct high-fidelity flow profiles with any desired T_s . We anticipate that the data generated by this self-built high resolution data collection system can help inform the selection of an optimal sampling interval for collection of ground-truth data for water end-use disaggregation purposes.

2) *Design Constraints*: We expect to deploy our data collection devices in a household setting. Water usage data from each outlet would be collected from our volunteers' home. Therefore, it is important to have the data collection devices run on battery, as it is not feasible to use mains electricity because of safety concerns and the lack of power sockets near water outlets. Moreover, the data collection devices to be deployed are expected to transmit all collected data autonomously in order to avoid the nuisance of dispatching crew to retrieve data manually on a regular basis. The ability to transmit data autonomously also enables us to analyze the collected data with reduced turnaround time. The design goal

is to power each device for multiple months on a single charge. This would enable us to study seasonal changes in domestic water consumption while reducing the frequency of battery replacement and hence disturbance to our participants.

C. System Design

We have designed and deployed a prototype IoT sensors system in 2020 for collecting tap-based domestic water consumption (i.e., the ground truth data for training of water end-use disaggregation models). Fig. 3 outlines the architecture of the data collection system. This section introduces four key building blocks contributing to the collection of high-resolution flow data: 1) Tap Sensor hardware; 2) Flow event logging algorithm, which runs on the Tap Sensor; 3) Household gateway; and 4) Cloud and on-premise back-end services.

1) *Tap Sensor Hardware*: A data collection device has been installed at each outlet of our volunteers' home to record the flow profile of each flow event of each water appliance. This device, which we refer to as a *Tap Sensor*, is composed of a Hall-effect turbine flow sensor and a custom-built, battery-powered IoT sensor node.

SIKA VTY10MA turbine flow sensor [25], a commercial off-the-shelf flow sensor, was used for measuring water flowing from each outlet. This flow sensor, in compliance with the local drinking water safety standards, has a threaded connection of DN10 diameter¹, which is suitable for in-line installation in residential plumbing pipes. It is constructed with a bipolar latching Hall-effect sensor. When water flows through the sensor body, its internal magnetized rotor rotates at a given rate. The Hall-effect sensor, integrated into the sensor body, latches its output voltage relative to the magnetic pole it senses. Continuous water flow then forces the turbine flow sensor to produce a pulse train of square waves. The pulse period is inversely proportional to flow rate. According to the datasheet, the VTY10MA turbine flow sensor is capable of sensing flow rates ranged from 1 to 30 L/min with a volumetric resolution of 2.02 mL per pulse.

The VTY10MA turbine flow sensor is interfaced with our custom-built sensor node, as shown in Fig. 4. The sensor node features a Microchip ATSAMD21G18 microcontroller (with a

¹Nominal pipe diameter of 10 mm, or 3/8 inch.

Participating Household

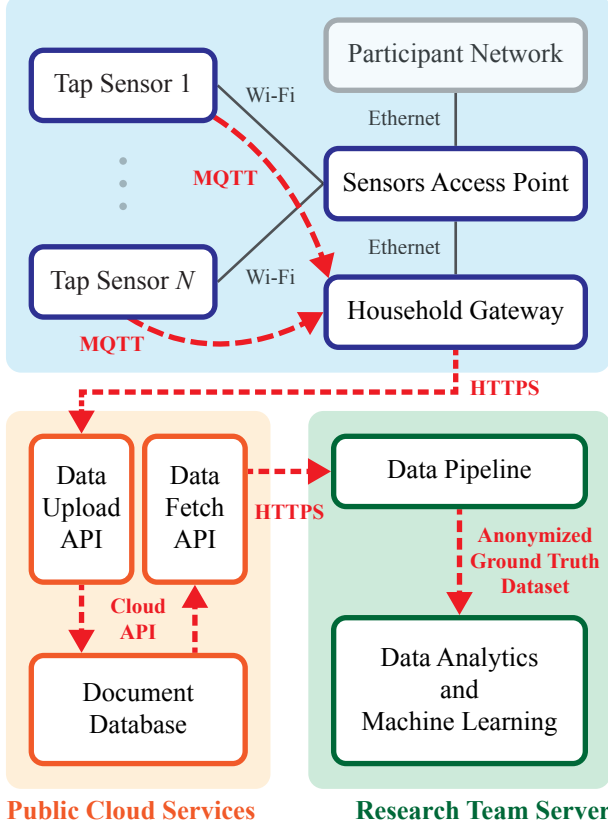


Fig. 3. System architecture of the proposed IoT sensors system. Dashed arrows indicate data pathway from Tap Sensors to the research team server.

48 MHz ARM Cortex-M0+ core and 32 KiB of SRAM) [26], an Espressif ESP8266 Wi-Fi communication module, and a 4 MiB non-volatile NOR flash memory chip. Fig. 5 presents the hardware block diagram of the Tap Sensor. When a flow event commences, the VTY10MA sensor pulses trigger an interrupt service routine of the microcontroller to put it into the measurement mode for recording of pulse periods. The microcontroller runs an algorithm to encode and compress the pulse periods, then stores the flow event to the NOR flash memory. The stored flow events are subsequently transmitted to a household gateway via Wi-Fi. The sensor node also generates and transmits an hourly health report, which includes its battery voltage and other device status metrics, for diagnosis and maintenance planning.

The Tap Sensor is designed to operate on battery power. We profiled its power consumption levels to determine the energy budget requirements. Fig. 6 shows the current trace of the Tap Sensor when a flow event takes place, and illustrates the variation of current between its three modes of operation: When there is no water flowing through the flow sensor, the sensor node remains in a low-power *sleep* mode. If a flow event begins, the rising edge from the flow sensor's first pulse triggers the *measurement* mode, where the microcontroller starts keeping track of the incoming pulses. After a flow event ends, the sensor node turns on its communication module to

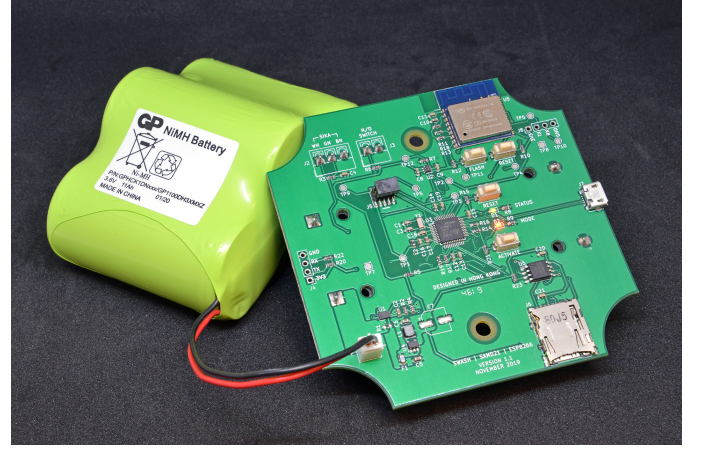


Fig. 4. Custom-built sensor node (right) with its NiMH battery pack (left).

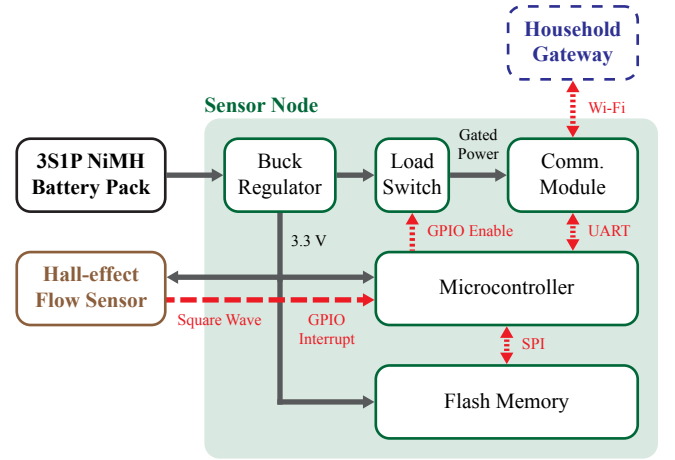


Fig. 5. Tap Sensor hardware block diagram showing the interconnection between sensor node components, the battery pack, and the Hall-effect turbine flow sensor. Solid and dashed arrows indicate DC power distribution and signal interface, respectively.

perform *Wi-Fi transmission* to send flow event data to the household gateway. Finally, it turns off the communication module and returns to the sleep mode until the next flow event occurs.

Although Wi-Fi transmission might appear to be the most *power-hungry* component in the current consumption profile, it is indeed not the major contributor to the overall *energy* consumption of the Tap Sensor. To collect ground truth data, we have to capture every flow event of each water outlet. In other words, the turbine flow sensor has to be powered around the clock so that not a spontaneous flow event is missed. The VTY10MA sensor, which consumes around 3 mA out of the 3.73 mA Tap Sensor sleep current, is the bottleneck of reducing energy consumption of our design. To keep the VTY10MA sensor in the hibernation mode in the absence of water flow, we have explored different low-power alternatives to detect the on-set of a flow event. Conceptually, there are viable options, such as the adding of an in-line reed switch to provide a binary indicator of flow, or a motion sensor to detect human presence around the water outlet. In practice, they are not robust options as a reed switch requires a relatively high

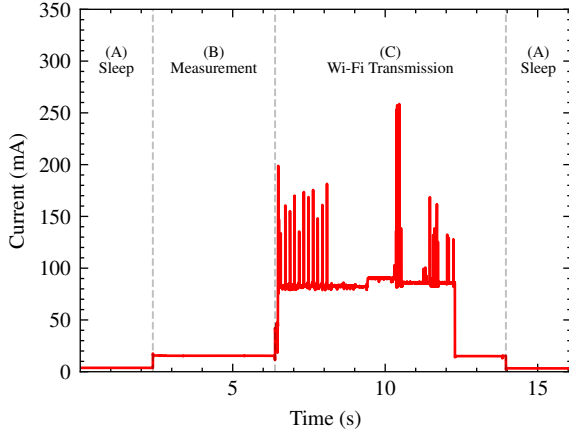


Fig. 6. Current consumption profile of the Tap Sensor in its three modes of operation: (A) sleep mode, with an average current of 3.73 mA; (B) measurement mode, with an average current of 15.5 mA; and (C) Wi-Fi transmission mode, with an average current of 85.2 mA. The current trace was captured by an Otii Arc source measurement unit.

flow rate to turn on. The high threshold for triggering on a reed switch would mean that low-flow-rate events would be left undetected. For motion sensors, they are unable to detect flow events generated by mechanized appliances (e.g., washing machine usage). Therefore, keeping the more sensitive Hall-effect turbine flow sensor powered on all the time remains our only feasible option.

Consequently, a high capacity battery is needed to sustain the Tap Sensor's operation for multiple months. Given the hot and humid environment in bathrooms and kitchens, we chose a nickel metal hydride (NiMH) rechargeable battery over lithium-ion-based batteries on safety grounds, in spite of the latter's higher energy density. To maximize lifespan, we used a pack of three 11 Ah, D-size NiMH cells connected in series (also known as a 3S1P configuration; shown in Fig. 4) to power each Tap Sensor. The sensor node and its battery pack are sealed in a waterproof box when deployed in the field.

2) *Flow Event Logging Algorithm*: Compared to existing flowmeters, the Tap Sensor has a distinctive capability of recording the finest details of a flow event such that a high-fidelity flow profile can be reconstructed with any given T_s . This is made possible by logging every pulse period r_k in the sequence of pulses generated by the turbine flow sensor. If there are N pulses generated during a flow event e_i , the sequence $S_i = \langle r_1, r_2, \dots, r_{N-1} \rangle$ holds every logged pulse period in milliseconds (ms). The duration of the flow event ($t_i^e - t_i^s$) can be approximated by $\sum_{k=1}^{N-1} r_k$. Assuming that each pulse corresponds to a fixed volume v flowed through the sensor, the total volume of e_i is $v \times N$.

While this portrays the train of pulses most accurately, recording S_i is memory-intensive and impractical on a constrained sensor node. To put this in perspective, a showering event that consumes 40 L of water would generate 19,800 pulses, given the VTY10MA flow sensor's nominal pulse rate of 495 pulses/L. Assuming that each r_k is represented as an unsigned 16-bit integer, S_i would take up 38.7 KiB, exceeding the

microcontroller's 32 KiB SRAM capacity. Hence, an encoding and compression scheme is needed to optimize memory usage. According to our empirical data, the VTY10MA produces r_k of very similar values under a steady flow rate. Thus, we can compress S_i using the principle of run-length encoding (RLE) [27]. Consider an original sequence of pulse periods $S_0 = \langle 23, 23, 24, 24, 46, 45, 45, 45 \rangle$. The simplest form of RLE algorithm compresses it into $\langle (23, 2), (24, 2), (46, 1), (45, 3) \rangle$. Each 2-tuple of (r_k, c_k) notates the number of consecutive occurrence(s) of r_k in S_0 as c_k . This may appear to be a more compact representation, but in fact it provides no memory usage improvement to S_0 : In this extreme example of a short event, both sequences contain 8 integer values. While most flow events have a longer duration and a larger value of N , hence a smaller overhead ratio, the simple RLE still causes some unnecessary overhead when there are minor fluctuations in a flow sensor's pulse output, and thus slight changes in r_k (e.g., 23 to 24 and 46 to 45).

We extended the RLE concept to devise a custom flow event logging algorithm, which groups very similar adjacent r_k values. The improved algorithm compresses S_i into $Y_i = \langle (\tau_1, c_1, \delta_1), (\tau_2, c_2, \delta_2), \dots, (\tau_M, c_M, \delta_M) \rangle$ where $M \leq N-1$. It has to run on-the-fly as the sensor node receives incoming pulses generated by the flow sensor, without the complete knowledge of S_i . As the second rising edge of e_i arrives at the sensor node, it marks the first pulse period as τ_1 . If a subsequent pulse period r_k satisfies $|r_k - \tau_1| < \lambda$ where λ is the maximum allowed deviation, c_1 increments and $(r_k - \tau_1)$ is added to δ_1 ; otherwise, a new 3-tuple is appended to Y_i , and the process repeats until the final pulse period is registered. The resultant flow event duration $D = \sum_{t=1}^M (\tau_t \times c_t + \delta_t)$. The pseudocode of the algorithm is listed in the box below:

```

Input: Sequence of pulse periods  $S = \langle r_1, r_2, \dots, r_{N-1} \rangle$ 
Output: Encoded and compressed sequence  $Y$ 
begin
  initialize an empty sequence  $Y$ ;
   $(\tau, c, \delta) \leftarrow (r_1, 1, 0)$ ;
  foreach  $r_k \in S \mid r_k \neq r_0$  do
    if  $|r_k - \tau| < \lambda$  then
       $c \leftarrow c + 1$ ;
       $\delta \leftarrow \delta + (r_k - \tau)$ ;
    else
      append  $(\tau, c, \delta)$  to  $Y$ ;
       $(\tau, c, \delta) \leftarrow (r_k, 1, 0)$ ;
    end
  end
  append  $(\tau, c, \delta)$  to  $Y$ ;
end

```

The constant λ is determined based on flow sensor characteristics. The VTY10MA sensor we used has a nominal pulse rate of 495 pulses/Litre, with a measurable flow rate of 1 to 30 L/min. Theoretically, these convert into a range of 8.25 to 247.5 pulses per second, or a pulse period range of 4 to 121 ms. As a trade-off between accuracy and memory usage, we set $\lambda = 3$ ms, meaning that all consecutive $r_k \in (\tau - 3, \tau + 3)$

are grouped. The maximum error due to compression would be less than 2.5% of the pulse period range. This error has a negligible distortion effect on actual flow rate estimation. With this algorithm, the example sequence S_0 is encoded and compressed into $Y_0 = \langle (23, 4, 2), (46, 4, -3) \rangle$.

Most machine learning models for water end-use disaggregation take a sequence of a flow event e_i 's volumetric usage $V_i = \langle v_1, v_2, \dots, v_j, \dots, v_P \rangle$ as input, where v_j is the volume at the j -th time step; each time step is spaced by a given T_s . The flow event logging algorithm facilitates flexible reconstruction of the required sequence: First, Y_i is converted into $S'_i = \langle r'_1, r'_2, \dots, r'_{N-1} \rangle$ by expanding and flattening each of the 3-tuples. (τ, c, δ) is expanded into c consecutive occurrence(s) of $\tau + (\delta/c)$. S'_i becomes a close approximation of S_i . Next, V_i is constructed as a sequence of length $\lceil D/T_s \rceil$. Then the value of each $v_j \in V_i$ is obtained by counting the number of pulses occurred within that time step, multiplied by the pulse rate. Flow profile p_i , which concerns the average flow rate instead of the volumetric usage, can be obtained through dividing each $v_j \in V_i$ by T_s .

3) *Household Gateway*: Our IoT sensors system adopts a client-gateway-server topology [28]. We deployed a *household gateway* and a dedicated Wi-Fi router as the access point for the sensors in each of the participating household, in order to support the installation of multiple Tap Sensors. The sensor access point connects to the Internet via the participating household's network. The household gateway is a Raspberry Pi single-board computer which runs customized software to serve two key functions in the data collection system.

First, it acts as a broker to enable reliable data exchange with the Tap Sensors using MQTT messaging protocol. The broker is set up with the *Eclipse Mosquitto* software. Authentication is required between the Tap Sensor and the gateway. The Tap Sensors publish compressed flow event data and hourly health reports, encoded as JSON objects, to the gateway. Due to limitations of the open-source MQTT client implementation on the Tap Sensor's communication module, QoS 1 (delivery at least once) and 2 (exactly once) [29] were not supported. We thus had to implement a custom acknowledgement (ACK) mechanism, to assure that every flow event and health report data object is successfully delivered to the gateway. If a Tap Sensor does not receive an ACK from the gateway, the transmission is considered unsuccessful. The Tap Sensor would then schedule a re-transmission at a later time, to ensure that all its logged flow events and health reports would be collected. Otherwise, the ACK signals the Tap Sensor to mark the data object as transmitted and to free up flash memory storage space. Also, the gateway keeps synchronized with a NTP time server, and publishes a Unix timestamp every second. This synchronizes the real-time clock of all Tap Sensors in the same household, ensuring that every flow event is tagged with a correct timestamp that is consistent throughout the household.

Second, the household gateway relays the Tap Sensor data to the cloud database. We consider the gateway-to-cloud segment as the least reliable link in the data collection system, due to various external factors, such as electricity and Internet service outage. Therefore, a resilient data transmission mechanism is in place to avoid missing data. The gateway subscribes to the

messages published by the Tap Sensors on specific MQTT message topics. Incoming data objects are placed on a *Redis* queue, cached in-memory and persisted to disk from time to time. Each of the queued data objects is scheduled for a secure transfer to the cloud server over HTTPS. Gateway metadata, including the timestamp received from the Tap Sensor, gateway identifier, and a hash of the content, are added to each data object. Upon a failed HTTPS request to the cloud server, the gateway would re-queue the data object for a later re-transmission, until the data object has been successfully sent at least once. In order to minimize unnecessary traffic, a timer backoff strategy, inspired by the TCP re-transmission timer for congestion control [30], has been implemented to limit the number of retries during network downtime.

There are a few design considerations in favor of this additional gateway layer, rather than connecting the Wi-Fi-enabled Tap Sensors directly to the cloud server. Cybersecurity is the primary consideration. By having a dedicated router and household gateway set-up, some prominent security risks in the IoT space can be mitigated. First, the router isolates the Tap Sensors network from the participating household's network, so that access control can be applied to the Tap Sensors. This also denies the Tap Sensors from accessing the Internet, and thus drastically reduces their attack surfaces. The risk of them being compromised by external bad actors to leak data, or exploited for launching DDoS attacks [31] can thus be mitigated. Second, the household gateway possesses substantially more processing power compared to the individual Tap Sensors, and is not constrained by battery power. More sophisticated encryption and hashing algorithms can be adopted to protect the confidentiality and integrity of the sensor data. In addition, the household gateway also enhances the energy efficiency of the Tap Sensors: The round-trip time for a Tap Sensor to communicate with the household gateway is significantly shorter than that of an Internet server. By establishing a connection to a known fixed IP within the same subnet, the time spent on DNS lookup and routing to the destination cloud server can be saved. This shortens each Wi-Fi transmission cycle, and hence, conserves energy and extends battery life.

4) *Back-end Services*: We have set up a secure cloud server to provide an authenticated RESTful API for the household gateways to upload flow event data and sensor health reports. The uploaded data objects are stored in a cloud-based document database. A data pipeline running on the research team's on-premise server download them in bulk periodically.

Due to intermittent network problems or occasional device failures, a Tap Sensor or a household gateway might transmit a data object more than once. The data pipeline is responsible for de-duplicating such re-transmitted data objects, by comparing the unique combination of a data object's timestamp, gateway identifier, and content hash. Anomalous flow events are also filtered in the data pipeline. These steps prepare a clean dataset for further analysis and machine learning.

In addition, Tap Sensor hardware and household gateway failure can be detected through each Tap Sensor's health reports. Each hourly health report serves as a "heartbeat" that tells the project team that a device is still alive. If the cloud

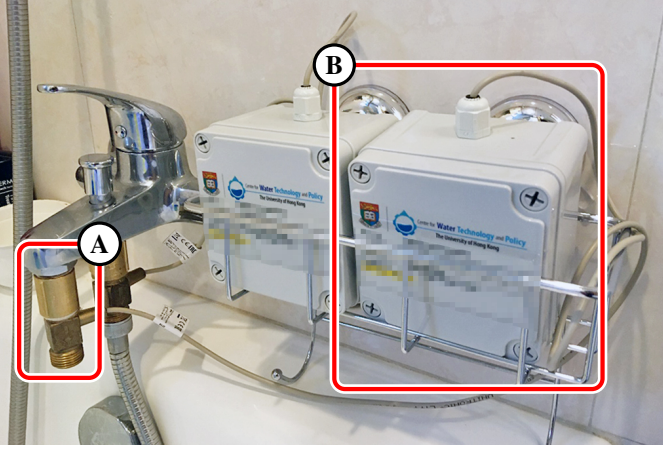


Fig. 7. An example of Tap Sensor installation at a bath-shower faucet. Each of the bathtub and the showerhead outlets is outfitted with (A) a Sika VTY10MA turbine flow sensor; and (B) a sensor node with battery enclosed in a waterproof box.

server does not register a particular Tap Sensor's health report for multiple hours, it is likely that the device has gone offline as a result of hardware failures, Wi-Fi connection drop-outs, or battery depletion. Similarly, missing reports from all Tap Sensors in a household would indicate an anomaly in the household gateway. During the field test, this health reporting mechanism provided valuable information for evaluating the performance of the Tap Sensors. It also facilitates prompt scheduling of maintenance visits to minimize downtime in water usage data collection.

IV. FIELD TEST

Custom-designed Tap Sensors introduced in Section III-C were rolled out to monitor water usage at all water outlets (including all bathtub taps, basin taps, kitchen taps, showerheads and washing machines) of three households in Hong Kong for a two-month field test between mid-March and mid-May 2020. These three households do not have irrigation systems and pools. Outdoor water-use is rare in Hong Kong, as most people live in high-rise apartment buildings. Dishwashers are also uncommon and these three households do not have a dishwasher. To support wireless, real-time data collection from the Tap Sensors, a household gateway and a dedicated Wi-Fi router were also installed in each apartment unit. An example of Tap Sensor installation is shown in Fig. 7. A total of 25 Tap Sensors were deployed across the three households.

The purpose of the field test is to thoroughly study the performance and reliability of the system in a domestic setting, such that we can identify and patch hardware and software issues before utilizing the system to collect domestic water consumption data for an extended period. In the following, we present the preliminary data collected, and discuss issues and challenges we encountered during the field test.

A. Flow Event Data Collection

All Tap Sensors recorded 27,285 distinct flow events across the three participating households during the two-month field

TABLE III
SUMMARY OF FLOW EVENTS COLLECTED DURING FIELD TEST

Parameter	Raw Dataset	Filtered Dataset
Number of flow events		
Total	27,285	20,258
Average per household per day	152	113
Duration of flow events ^a		
Total	125.70 h	118.79 h
Average	16,585 ms	21,110 ms
Median	5297 ms	7358 ms
Minimum	2 ms	76 ms
Average per household per day	2514 s	2376 s

^a Due to memory constraints, the maximum duration of each flow event recorded by a Tap Sensor is limited to 30 minutes. An event longer than that, which is rare in domestic water-use, is split into multiple events.

test period. Table III summarizes several parameters of the collected flow event data.

When we analyzed raw data collected, we noticed anomalous flow events characterized by unrealistically high flow rates, or extremely-low or extremely-high level of water-use, which could hardly be possible in a domestic setting. These events have $\tau < 4$ ms (i.e., the theoretical lower bound of the VTY10MA's pulse period), or have high variability in τ . In actual water-use events, τ tends to remain stable and exhibits only gradual change patterns. Frequent fluctuation of flow rates is unlikely. On the other hand, a noisy event would produce a signature of high variability and randomized τ . Noisy events are possibly results of pressure fluctuations in the plumbing system or electrical noise originated from its Hall-effect element; and these mechanical and electronic noise were picked up by the turbine flow sensor.

To remove noisy events from the dataset, we have developed a filter based on median absolute deviation (MAD) of a flow event's pulse periods. The MAD of an event is given by $\text{median}(|r'_k - \text{median}(r'_k)|) \forall r'_k \in S'_i$. Typically, an event involving actual water usage yields a very low MAD due to stable flow rate and hence r_k . On the contrary, based on our empirical data, a noisy event that leads to a false but substantial volume of water usage usually has an MAD of over 500 ms. The result of MAD filtering is visualized in Fig. 8(b), in comparison to the raw, unfiltered dataset shown in Fig. 8(a). MAD filtering eliminates invalid flow events that would contaminate the dataset for machine learning training purposes. This approach also demonstrates better efficacy than a naive filter based on the flow event duration or pulse count, as the latter requires a threshold that cannot be determined objectively. A poorly determined threshold might remove some valid flow events with short duration. Moreover, Fig. 8(c) and (d) illustrate that some anomalous flow events (with low volume but unusually long duration) remained in the dataset.

B. Tap Sensor Battery Life

Fig. 9 compares the battery discharge curves of six deployed Tap Sensors in one of the households. Their lifespans ranged

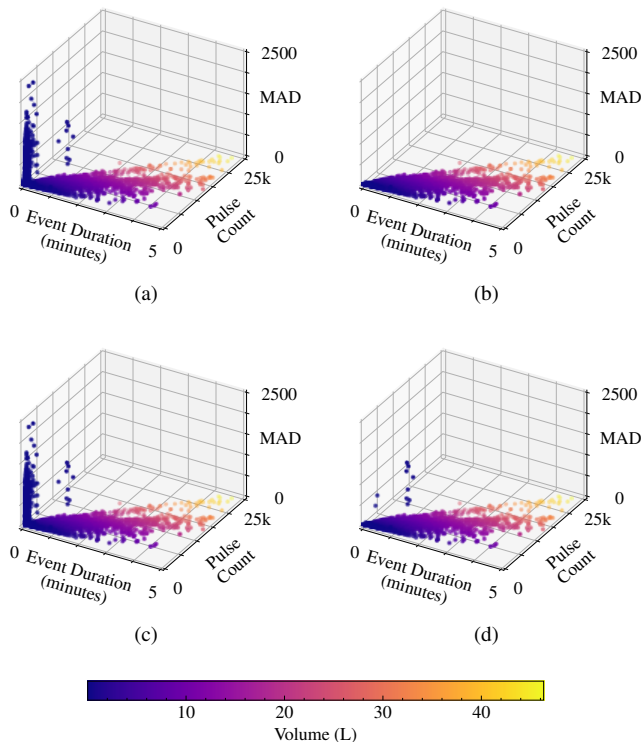


Fig. 8. Comparison of noisy flow event filtering results. (a) No filtering. (b) MAD filtering. (c) Event duration filtering. (d) Pulse count filtering. Each dot in the scatter plot represents a flow event with its volume color-coded.

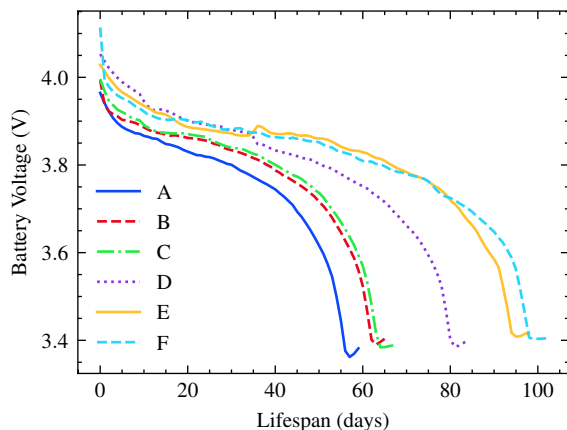


Fig. 9. Battery discharge curves of six Tap Sensors. They were deployed to the same household, but subjected to varying total flow duration and Wi-Fi signal quality at different water outlets.

from 58 (sensor A) to 101 (sensor F) days² of continuous operation. The variation in lifespan is primarily due to differences in 1) the total flow duration a Tap Sensor recorded; and 2) Wi-Fi signal quality. Table IV lists the factors affecting their lifespans, namely total flow duration, flow events count and median Wi-Fi connection time.

²Some Tap Sensors started operating before the 60-day field test period started; and some continued to operate after the end date. The six selected devices' lifespans and flow event statistics were collected after complete depletion of their batteries.

TABLE IV
FACTORS AFFECTING TAP SENSOR BATTERY LIFE

Tap Sensor	Number of Flow Events	Total Flow Duration (s)	Median Wi-Fi Conn. Time (s)	Lifespan (days)
A	114	2232	6.435	58
B	167	6282	5.084	64
C	1491	23,625	5.067	66
D	233	4115	5.007	83
E	159	773	6.633	97
F	34	922	5.550	101

The total flow duration a Tap Sensor has endured determines how much time it has spent in the measurement mode. Section III-C described that the average current consumption of flow measurement is around five times higher than sleep mode. Tap Sensors set up at more frequently-used water outlets, such as sensors B and C, had shorter lifespans than the rarely-used sensors E and F. In the initial firmware design, a Wi-Fi transmission takes place after each flow event is recorded for real-time diagnostics; therefore, having more flow events recorded would have impacted the battery life. Moreover, weaker Wi-Fi signal lengthens the transmission cycle, and significantly increases the energy usage due to the high current consumption during transmission. The median connection time in Table IV is a proxy indicator of a Tap Sensor's signal quality at different locations within the household. The better the wireless connectivity, the faster the connection can be established, and vice versa. This suggests that sensor A experienced the worst Wi-Fi signal quality, and hence the shortest lifespan among the six.

C. System Reliability

The resilience protocol outlined in Section III-C was demonstrated during the field test. Fig. 10 presents the distribution of transmission latency in all received flow events. Transmission latency is defined as the time difference between the end of a flow event and the event's insertion to the cloud database. A vast majority of flow events were successfully transmitted immediately after their occurrence. There were 1.30 % of flow events transmitted one hour or later after occurrence, and 0.42 % one day or later after flow event occurrence. The ACK and re-transmission mechanisms contributed to a high reliability of data delivery.

Each Tap Sensor keeps track of the number of flow events it has recorded onboard, and transmits this accumulated number in its hourly health report. We have compared the number of flow events reported by the device with the actual number of flow events received by the cloud database. Only five out of 27,285 flow events were not transmitted successfully. In other words, the overall successful delivery ratio is 99.98 %.

D. Deployment Issues

We rolled out the Tap Sensors household by household and streamlined the installation and maintenance procedures bit by

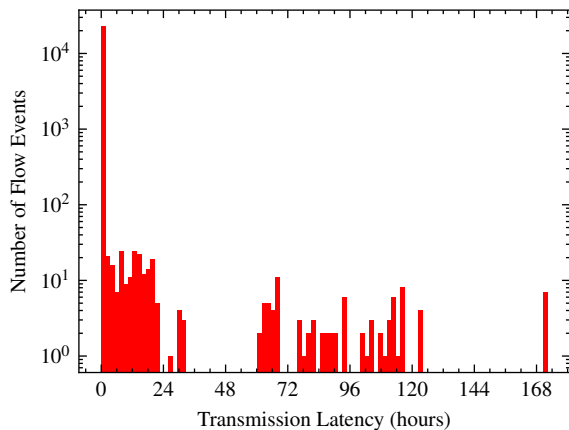


Fig. 10. Distribution of flow event transmission latency. The number of flow events axis is drawn in logarithmic scale.

bit. This gradual, incremental strategy allowed us to spot and rectify two systemic issues early on. First, a couple of Tap Sensors failed to report any flow event a few days after installation, while they continued to transmit health reports. The account of water consumption activities provided by our participants, nevertheless, had rejected the null usage hypothesis. It was revealed later that the rotor of the turbine flow sensors of these Tap Sensors were clogged by impurities built up inside the pipe. The clogs prevented the flow sensors from generating pulses. We mitigated this unexpected mechanical failure mode by installing a screen filter before the flow sensor. Second, we found a sensor node suffered from water damage, due to improper installation of the waterproof box accessories. Since then, we improved our on-site sealing and checking procedures to avoid similar human errors.

Nevertheless, there were still unexpected wireless connectivity issues. Radio signals were attenuated after penetrating multiple walls to reach bathroom basin and kitchen cabinets. It is difficult to maintain stable connections to some Tap Sensors with degraded Wi-Fi signal quality. The intermittent drop-outs of these Tap Sensors had not only interrupted data transmission, but also reduced their lifespans. We managed to mitigate this problem by installing Wi-Fi repeaters in the affected households. However, the level of commitment required of the participants increased further if Wi-Fi repeaters have to be installed, as participants have to free up an additional power socket and provide power to have Wi-Fi repeaters set up.

Last but not least, we reckoned that battery life expectancy remains the key challenge of deploying tap-based sensors to household environments. Even though our field tests have demonstrated that Tap Sensor E could last for over 100 days, it should be noted that the frequency of battery replacement for an entire household is indeed dictated by the shortest-lived device, i.e., around two months. The operational costs of deploying Tap Sensors to monitor household water consumption activities could only be reduced if the battery life expectancy of the shortest-lived device is extended.

V. RESULTS AND DISCUSSION

The field test proves that our IoT sensors system can record high-resolution domestic water consumption data and transmits them for analysis in near real-time. This unique capability enabled the collection of tap-based ground truth data in actual household environment for machine learning purposes. After concluding the two-month field test, we inspected the deployed Tap Sensors thoroughly, and made minor changes to their hardware and software, in order to rectify issues discussed in Section IV and to further enhance system reliability. After making sure that all the Tap Sensors were operating normally, the same data collection set-up (comprising the Tap Sensors, gateway infrastructure, as well as a secure data processing pipeline described in Section III) was used to collect water consumption data from the three participating households for another six months. In total, we gathered 61,923 flow events over a span of eight months (which comprises a two-month field test period and a six-month data collection period). In this section, we use our empirical data to investigate the relationship between different sampling intervals T_s and the performance of machine learning algorithms for water end-use classification.

Autoflow v3.1 [8] is the closest and the most up-to-date reference, in the domain of water end-use classification. It applied the Hidden Markov model (HMM) as one of its major classifiers in its disaggregation pipeline [8]. Support-vector machine (SVM) has been widely used by scholars, such as [4], [24], to solve water disaggregation problems. Therefore, we used our ground truth data to examine the classification accuracy of SVM and HMM models. Datasets used to train SVM and HMM models were constructed using different T_s . It is expected that a shorter T_s would yield a higher accuracy as we hypothesized in Section III-A, regardless of the type of machine learning algorithms.

SVM requires a set of features as input for the classifier. The choice of features affects its classification performance. All the flow rate-related parameters, as listed in Table II, were included in our feature set for analysis. Additionally, we have included the hour of the day of flow event occurrence as [4] suggested. Since we have only collected tap-based water consumption data, we did not include supplementary features explored in the literature like electricity and hot water usage. HMM, on the other hand, takes flow profiles as input. We prepared the volumetric usage sequence V_i of each flow event e_i sampled at different T_s , using the procedures described in Section III-C.

Filtered tap-based water consumption data from all participating households were combined to a single dataset, which is then split into a training set and a test set with the ratio of 80/20. In this way, a household-agnostic *universal classifier* can be built. Machine learning models were trained to classify flow events into four water end-use categories, namely basin use, kitchen use, showering and use by washing machine. Basin use and kitchen use refer to usage takes place at the bathroom basin taps and kitchen taps, respectively. Showering refers to water-use events recorded by both the showerheads and the bathtub taps.

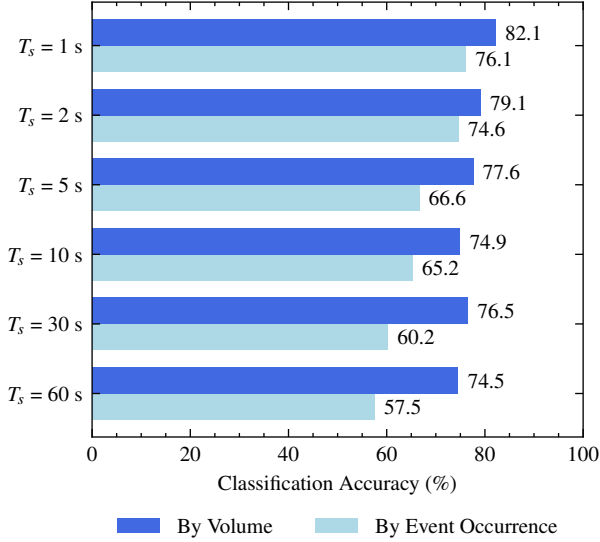


Fig. 11. Effect of sampling interval T_s on the classification accuracy of the SVM-based universal classifier. Both accuracy by volume and by event occurrence decrease with an increasing T_s .

Fig. 11 presents the classification accuracy of our SVM model trained with datasets of $T_s = 1$ s, 2 s, 5 s, 10 s, 30 s, and 60 s. *Classification accuracy by event occurrence* refers to the percentage of correctly classified events in the test set. *Classification accuracy by volume* refers to the ratio between “total volume of water-use in correctly classified events” and “total volume of water-use in all flow events” [24]. It is clear that the longer the T_s , the lower the level of classification accuracy. The accuracy by event occurrence was reduced from 76.1 to 57.5 % as T_s increased from 1 to 60 s. In particular, it suffered a significant loss of 8 % when T_s increased from 2 to 5 s. On the other hand, the accuracy in terms of prediction of volume of use by different end-use categories is higher than that of event labeling. Classification accuracy by volume has a relatively less severe drop as well (from 82.1 to 74.5 %) as T_s increased from 1 to 60 s. By referring to the confusion matrices, it is found that the differences of drop of accuracy against T_s is a result of the distinctive use pattern of the kitchen taps and the basin taps. Kitchen use and basin use were characterized by a large number of short-duration flow events, which are more prone to flow profile distortion with long T_s . On the contrary, the long-duration showering and washing machine events are less susceptible to the impacts of long T_s . As long-duration showering and washing machine events contributed to around 70 % of total freshwater use, classification accuracy by volume tends to be higher even with longer T_s .

We attempted to build a HMM-based universal classifier, but unfortunately, our training dataset of limited number of households could only yield an accuracy level of less than 50 %. In order to study the effect of T_s on SVM- and HMM-based models, we have built a household-specific classification model for one of the households instead. Using the household’s water consumption data as training data, the household-specific classification model could perform end-use

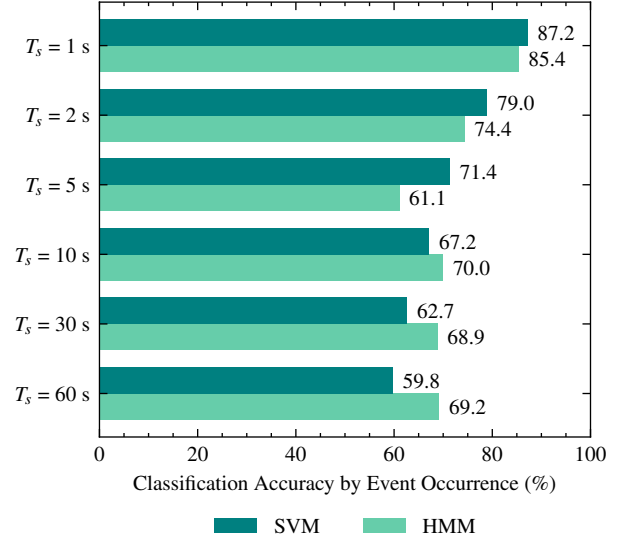


Fig. 12. Effect of sampling interval T_s on the classification accuracy by event occurrence of SVM- and HMM-based household-specific classifiers. Accuracy generally decreases with a longer T_s , regardless of the algorithm type.

classification for future flow events generated in the same household. Fig. 12 presents the classification accuracy by event occurrence of our SVM- and HMM-based household-specific classifiers. The accuracy of SVM deteriorated from 87.2 to 59.8 % as T_s increased from 1 to 60 s. The accuracy of HMM exhibited a downward trend in general, with its accuracy level falling from 85.4 to 69.2 % as T_s increases from 1 to 60 s. However, there is no plausible explanation for the dip in accuracy at $T_s = 5$ s.

To summarize, we proved that a high-resolution dataset can boost the accuracy level of water end-use classification. Performance of classifiers would benefit from a dataset with shorter sampling interval, regardless of the machine learning algorithms used. We conclude that high-resolution data with 1-second sampling interval offer the best classification performance. However, most of the market-available flowmeters or AMR systems are unable to generate such fine-grained data. Based on our preliminary results, $T_s = 5$ s may be a reasonable trade-off, as it still provides satisfactory classification accuracy by volume especially when SVM-based models are used. Further increase in T_s would hamper the level of classification accuracy significantly.

VI. CONCLUSION

The use of fine-grained tap-based consumption data as ground truth data can benefit the development of an accurate end-use disaggregation algorithm, which is essential for conducting *Smart Water Auditing*. Understanding how much water has been used and for what purposes build the fundamental ground for developing effective water conservation policy. This work has contributed to an high-resolution tap-based IoT system for collecting ultra-high-resolution tap-based water use data that has never been collected before. The collected data is essential for training accurate water disaggregation water machine learning model for water end-use analysis. It

also allows the research team to systematically investigate the impacts of different sampling intervals on the accuracy level of the water end-use disaggregation model.

To collect fine-grained tap-based consumption data from multiple households, we have built a set of Wi-Fi-based Tap Sensors, gateway infrastructure, as well as a secure data processing pipeline. A compression mechanism was devised to allow our Tap Sensors to record ultra-high-resolution water flow data. This first-of-its-kind data collection system has not only gathered the necessary data for machine learning, but also enabled a systematic investigation into the impacts of different sampling intervals on the accuracy level of water end-use disaggregation models. The purpose-built system was commissioned successfully to 25 water outlets from three households. The 8-month field test and data collection exercise has also validated that our Tap Sensors are resilient by design, and they are highly reliable in terms of the ability to transmit data in challenged wireless network environments. Over 60,000 tap-based flow events were used to train a preliminary SVM-based end-use classifier. An accuracy level of 82.1 % was attained. We have also proved that very fine-grained water flow data (with sampling interval at 1 second) can boost the performance of both SVM and HMM models substantially. Given that the primary goal of water auditing is to understand the composition of water use in the domestic sector, efforts should be put to further improve accuracy of models in predicting the volume of use by different end-use categories.

To collect ground truth data for the building of a generalizable water end-use disaggregation model, we plan to scale up the data collection exercise to have more Tap Sensors installed in more households. Wireless connectivity and energy consumption are the two major barriers to a scalable and robust data collection system. In order to eliminate the need of setting up an additional network in participating households and to reduce energy usage of Tap Sensors incurred by data transmission, low-power wide-area network (LPWAN) technologies will be explored. Minimizing energy usage for flow measurement is even more critical, as the need to provide energy for powering and supporting the power-hungry Hall-effect turbine flow sensors has been a bottleneck for battery life extension. We will research into and develop techniques for reducing the average power consumption by the turbine flow sensor through adaptive duty cycle mechanisms, such that a Tap Sensor can last for multiple months using a smaller battery pack. The use of LPWAN technologies and a more compact Tap Sensor will help minimize the level of disturbances brought to participants in the future.

ACKNOWLEDGMENT

The authors thank Dr. Benjamin Ho for his assistance in the data collection exercise.

REFERENCES

- [1] F. Lee and A. Lee, "Smart technologies, actionable data, and domestic water consumption in Hong Kong: Potentials and constraints," International Water Resources Association, Tech. Rep., 2021, pp. 160–169. [Online]. Available: <https://www.iwra.org/wp-content/uploads/2022/02/Rapport-complet-web-ok-1.pdf>.
- [2] J. Sturman, G. Ho, and K. Mathew, *Water Auditing and Water Conservation*. IWA Publishing, 2005, ISBN: 9781780402710. DOI: 10.2166/9781780402710.
- [3] J. O. Nelson, "Water Audit Encourages Residents to Reduce Consumption," *Journal - American Water Works Association*, vol. 84, no. 10, pp. 59–64, 1992, ISSN: 1551-8833. DOI: 10.1002/j.1551-8833.1992.tb05866.x.
- [4] Y. Gao, D. Hou, N. K. Banerjee, and S. Banerjee, "Water Fixture Identification in Smart Housing: A Domain Knowledge Based Case Study," *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 904–909, 2016. DOI: 10.1109/icmla.2016.0162.
- [5] L. Pastor-Jabaloyes, F. J. Arregui, and R. Cobacho, "Water End Use Disaggregation Based on Soft Computing Techniques," *Water*, vol. 10, no. 1, p. 46, 2018. DOI: 10.3390/w10010046.
- [6] M. Wonders, Z. Ghassemlooy, and M. A. Hossain, "Training with synthesised data for disaggregated event classification at the water meter," *Expert Systems with Applications*, vol. 43, pp. 15–22, 2016, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2015.08.033.
- [7] K. A. Nguyen, R. A. Stewart, H. Zhang, and C. Jones, "Intelligent autonomous system for residential water end use classification: Autoflow," *Applied Soft Computing*, vol. 31, pp. 118–131, 2015, ISSN: 1568-4946. DOI: 10.1016/j.asoc.2015.03.007.
- [8] A. Yang, H. Zhang, R. A. Stewart, and K. Nguyen, "Enhancing Residential Water End Use Pattern Recognition Accuracy Using Self-Organizing Maps and K-Means Clustering Techniques: Autoflow v3.1," *Water*, vol. 10, no. 9, p. 1221, 2018. DOI: 10.3390/w10091221.
- [9] K. A. Nguyen, R. A. Stewart, H. Zhang, and O. Sahin, "An adaptive model for the autonomous monitoring and management of water end use," *Smart Water*, vol. 3, no. 1, pp. 1–21, 2018.
- [10] C. Beal and R. A. Stewart, "South East Queensland Residential End Use Study: Final Report," Tech. Rep., 2011.
- [11] A. L. Karn, S. Pandya, A. Mehbodniya, F. Arslan, D. K. Sharma, K. Phasinam, M. N. Aftab, R. Rajan, R. K. Bommisetti, and S. Sengan, "An integrated approach for sustainable development of wastewater treatment and management system using IoT in smart cities," *Soft Computing*, 2021, ISSN: 1433-7479. DOI: 10.1007/s00500-021-06244-9. [Online]. Available: <https://doi.org/10.1007/s00500-021-06244-9>.
- [12] M. S. Rahim, K. A. Nguyen, R. A. Stewart, D. Giurco, and M. Blumenstein, "Advanced household profiling using digital water meters," *Journal of Environmental Management*, vol. 288, p. 112377, 2021, ISSN: 0301-4797. DOI: <https://doi.org/10.1016/j.jenvman.2021.112377>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0301479721004394>.
- [13] L. Yang, S.-H. Yang, E. Magiera, W. Froelich, T. Jach, and C. Laspidou, "Domestic water consumption

- monitoring and behaviour intervention by employing the internet of things technologies,” *Procedia Computer Science*, vol. 111, pp. 367–375, 2017, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2017.06.036>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050917312097>.
- [14] R. Khan, S. U. Khan, R. Zaheer, and S. Khan, “Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges,” *2012 10th International Conference on Frontiers of Information Technology*, pp. 257–260, 2012. DOI: 10.1109/fit.2012.53.
- [15] A. Cominola, M. Giuliani, D. Piga, A. Castelletti, and A. Rizzoli, “Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review,” *Environmental Modelling & Software*, vol. 72, pp. 198–214, 2015, ISSN: 1364-8152. DOI: 10.1016/j.envsoft.2015.07.012.
- [16] A. Roshan and M. Kumar, “Water end-use estimation can support the urban water crisis management: A critical review,” *Journal of Environmental Management*, vol. 268, p. 110663, 2020, ISSN: 0301-4797. DOI: 10.1016/j.jenvman.2020.110663.
- [17] W. B. DeOreo, P. Mayer, B. Dziegielewski, and J. Kiefer, *Residential End Uses of Water, Version 2: Executive Report*, 2016.
- [18] A. Liu, D. Giurco, and P. Mukheibir, “Urban water conservation through customised water and end-use information,” *Journal of Cleaner Production*, vol. 112, pp. 3164–3175, 2016, ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2015.10.002.
- [19] *Albuquerque Single Family Water Use Efficiency and Retrofit Study*, Dec. 2011. [Online]. Available: <https://aquacraft.com/wp-content/uploads/2016/07/Albuquerque-Report-Exec-Summary.pdf> (visited on 12/01/2020).
- [20] *California Single-Family Water Use Efficiency Study*, Dec. 2011. [Online]. Available: <https://aquacraft.com/wp-content/uploads/2015/10/CalSF-Executive-Summary.pdf> (visited on 12/01/2020).
- [21] S. Gato-Trinidad, N. Jayasuriya, and P. Roberts, “Understanding urban residential end uses of water,” *Water Science and Technology*, vol. 64, no. 1, pp. 36–42, 2011, ISSN: 0273-1223. DOI: 10.2166/wst.2011.436.
- [22] R. Wills, R. Stewart, and B. Capati, “Closing the loop on water planning: an integrated smart metering and web-based knowledge management system approach,” in *10th IWA International Conference on Instrumentation Control and Automation*, 2009. [Online]. Available: <http://hdl.handle.net/10072/31802>.
- [23] K. A. Nguyen, H. Zhang, and R. A. Stewart, “Development of an intelligent model to categorise residential water end use events,” *Journal of Hydro-environment Research*, vol. 7, no. 3, pp. 182–201, 2013, ISSN: 1570-6443. DOI: 10.1016/j.jher.2013.02.004.
- [24] J. Vitter and M. Webber, “A non-intrusive approach for classifying residential water events using coincident electricity data,” *Environmental Modelling & Software*, vol. 100, pp. 302–313, 2018, ISSN: 1364-8152. DOI: 10.1016/j.envsoft.2017.11.029.
- [25] Sika Systemtechnik GmbH, *Turbine Flow Sensor Series VTY [Operating Manual]*, 2020. [Online]. Available: https://www.sika.net/fileadmin/products-multi/user_manuals/turbine_flow_sensors/Ea7900_VTY.pdf.
- [26] Microchip Technology Inc., *SAM D21/DA1 Family [Complete Datasheet]*, 2021. [Online]. Available: <https://ww1.microchip.com/downloads/en/DeviceDoc/SAM-D21DA1-Family-Data-Sheet-DS40001882G.pdf>.
- [27] A. Apostolico, G. Landau, and S. Skiena, “Matching for run-length encoded strings,” *Proceedings. Compression and Complexity of SEQUENCES 1997 (Cat. No.97TB100171)*, pp. 348–356, 1997. DOI: 10.1109/sequen.1997.666929.
- [28] C.-W. Yau, T. T.-O. Kwok, C.-U. Lei, and Y.-K. Kwok, “Energy Harvesting in Internet of Things,” in *Internet of Everything*, Singapore: Springer, 2018, pp. 35–79, ISBN: 978-981-10-5861-5. DOI: 10.1007/978-981-10-5861-5_3.
- [29] N. Naik, “Choice of Effective Messaging Protocols for IoT Systems: MQTT, CoAP, AMQP and HTTP,” *2017 IEEE International Systems Engineering Symposium (ISSE)*, pp. 1–7, 2017. DOI: 10.1109/syseng.2017.8088251.
- [30] V. Jacobson, “Congestion avoidance and control,” *ACM SIGCOMM Computer Communication Review*, vol. 18, no. 4, pp. 314–329, 1988, ISSN: 0146-4833. DOI: 10.1145/52324.52356.
- [31] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, “DDoS in the IoT: Mirai and Other Botnets,” *Computer*, vol. 50, no. 7, pp. 80–84, 2017, ISSN: 0018-9162. DOI: 10.1109/mc.2017.201.