Demo Abstract: Fault Diagnosis System for Low-cost Air Pollution Sensors

Sumukh Marathe, Akshay Nambi, Nishant Shrivastava, Manohar Swaminathan, Ronak Sutaria*
Microsoft Research India, *Respirer Living Sciences Pvt. Ltd
akshayn@microsoft.com

ABSTRACT

Fine-grained air pollution monitoring is a fundamental step towards curbing pollution levels. This is sought to be achieved by the large-scale deployment of low-cost sensors at high spatio-temporal resolution. Due to the nature of these deployments, in-the-wild and in harsh environments, sensors are prone to failures and hence ensuring data reliability is challenging. Furthermore, detecting a fault by analyzing the sensor data using existing data-centric approaches is non-trivial. This demonstration presents a sensor fault diagnosis system that employs the current signature of the sensor to address data reliability issues. The current signature captures the electrical characteristics of the hardware components enabling accurate detection and isolation of faults in low-cost pollution sensors.

CCS CONCEPTS

• **Computer systems organization** → **Embedded and cyber-physical systems**; • **Hardware** → **Fault tolerance**.

KEYWORDS

Fault detection and isolation; Air Pollution, PM 2.5 faults

1 INTRODUCTION

Air pollution is a major concern worldwide, with an estimated 7 million deaths every year [6]. Pollutants such as particulate matter of 2.5 microns or less, i.e., $PM_{2.5}$ is a major factor contributing to this mortality [6]. Hence, $PM_{2.5}$ monitoring at a fine-grained resolution is key to identify potential sources and raise awareness.

Traditional pollution monitoring systems rely on accurate, and expensive sensors, with each costing over $20,000. These are run by the government institutions in a few sparse locations. However, air pollution is known to be a complex phenomenon with spatio-temporal variations requiring fine-grained, and hyper-local measurements. With advances in sensing technologies, recent efforts have employed low-cost sensors for fine-grained $PM_{2.5}$ monitoring at scale. These sensors are compact, portable, and typically cost between $30 to $100 [1, 2]. The typical operation of a low-cost sensor is based on light scattering principle, where a small DC FAN controls the airflow allowing the particles to pass through a beam of light (usually emitted by an LED). Light is scattered by the particles, which is then detected by a photodiode and converted into particle count and mass concentration values (in $\mu g/m^3$).

SenSys '20, November 16–19, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-7590-0/20/11.
https://doi.org/10.1145/3384419.3431191

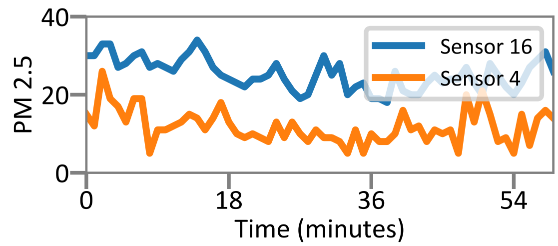


Figure 1: **Sensor data faults, Sensor 16 (working) and Sensor 4 (faulty).**

Recent works report several challenges in sensor reliability leading to significant data inaccuracies [4, 5]. Data inaccuracies are mainly associated with faults in low-cost sensors and its components [4]. The key components of $PM_{2.5}$ sensor include, the FAN and LED, where faults may arise due to the LED no longer emitting light and/or FAN no longer rotating at the correct speed. For example, if the LED stops working (i.e., not emitting light) then the $PM_{2.5}$ sensor will continuously read low values as there is no light scattered and is non-trivial to detect without sensor redundancy.

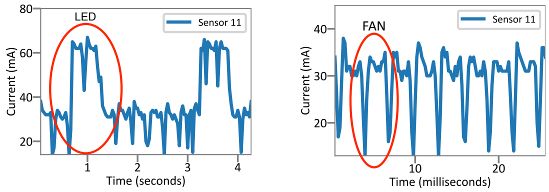
Prevalent research has focused on analyzing sensor data to identify fault patterns by detecting anomalies [5]. However, given the hyper-local variations in air pollution, it is extremely challenging to detect sensor faults, especially when a faulty sensor mimics a working sensor [4, 5] and anomalous data need not represent a faulty sensor. This issue is highlighted in Figure 1, where *a faulty sensor data mimics a working sensor data*. Both, Sensor 16 and Sensor 4 were deployed in different outdoor locations and it can be seen that the data from both the sensors are in similar range. However, upon manual inspection, Sensor 4 had a FAN fault leading to non-uniform airflow resulting in inaccurate PM data.

2 FAULT DIAGNOSIS SYSTEM FOR $PM_{2.5}$ SENSORS

We now describe our novel approach to reliably detect, and isolate sensor faults, going beyond traditional data-centric approaches.

Given the understanding of the working of low-cost $PM_{2.5}$ sensors, *our hypothesis is that all the data faults encountered in $PM_{2.5}$ sensors can be mapped to a failure or degradation of hardware components such as the FAN and LED* [4]. Typical, $PM_{2.5}$ sensors are compact digital sensors and hence, it is impossible to monitor any internal signals of the sensor components, that could have provided insight on the hardware characteristics.

Thus, to determine if a sensor is working or faulty, we measure *the current drawn by the sensor*, which captures the electrical characteristics of all the hardware components present in a sensor. The intuition here is that when a sensor goes faulty it's hardware characteristics vary, leading to changes in its current draw. We need to monitor only the overall current drawn by the sensor to determine its status. Hence it is non-intrusive, i.e., does not require breaking open the sensor to monitor the signals and requires no hardware



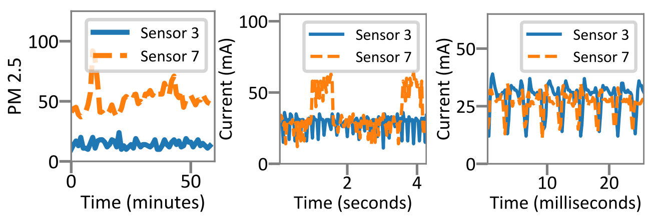
(a) Current @ 30Hz. **(b) Current @ 5kHz.**

Figure 2: Current signature sampled at 30Hz & 5kHz for a working sensor.

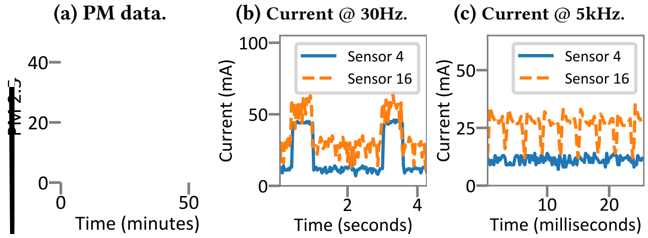
modification. This current measurement can be done using a current monitoring circuit (i.e., current sense amplifiers [3]), which are commonly present in battery-powered low-cost devices.

A $PM_{2.5}$ sensor includes a set of electro-mechanical components each operating at different frequencies. For instance, in a popular $PM_{2.5}$ sensor from Plantower [2], the LED switching on/off frequency is 0.4Hz (i.e., every 2.5s) and the FAN spins at around 200Hz (12000 RPM) (derived from the datasheet [2]). Thus, in order to capture all the variations, we sample current at both low and high frequencies. Specifically, in the case of Plantower sensor, we sample at 30Hz and 5kHz to monitor both low-frequency (LED) and high-frequency components (FAN) of the sensor. The selection of sampling rate is guided by information from sensor datasheet and varies from one manufacturer to another.

Figure 2(a) and (b) show the current drawn by a working Plantower $PM_{2.5}$ sensor sampled at 30Hz and 5kHz. The highlighted region indicates the change in current drawn due to the LED and FAN components turning ON, respectively.



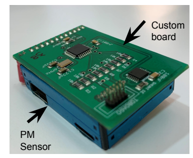
(a) PM data. **(b) Current @ 30Hz.** **(c) Current @ 5kHz.**

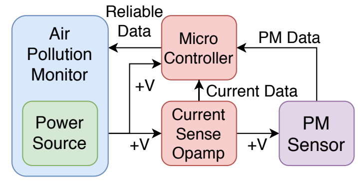


(d) PM data. **(e) Current @ 30Hz.** **(f) Current @ 5kHz.**

Figure 3: Fault detection and isolation in $PM_{2.5}$ sensors.

We now highlight how sensor current signature can be used to detect and isolate faults in a $PM_{2.5}$ sensor. Figure 3(a) shows data from two co-located sensors, Sensor 7 (working sensor) and Sensor 3 (faulty sensor with a LED malfunction). As we can see the data from both the sensors are in a similar range and is quite difficult to detect the faulty sensor without sensor redundancy. However, with the help of the current signatures, we can accurately detect and isolate the faults, if any, in these sensors. Figure 3(b) shows the current signature sampled at 30Hz for both the sensors and we can clearly see the distinct current signature of a faulty LED component of Sensor 3 as compared to Sensor 7. Furthermore, the





(a) **(b)**

Figure 4: (a) Plantower $PM_{2.5}$ sensor retrofitted with custom board (b) Block diagram of air pollution monitoring device used in the demo.

current signature sampled at 5kHz (Figure 3(c)) looks identical for both the sensors, indicating no fault in the FAN component.

Similarly, Figure 3(d) shows data from two co-located sensors - Sensor 16 (working sensor) and Sensor 4 (faulty with a FAN malfunction). Figures 3(f) shows distinct current signature sampled at 5kHz for Sensor 4 as compared to Sensor 16, indicating a FAN fault in Sensor 4. Furthermore, the current signature sampled at 30Hz (Figure 3(e)) looks identical for both the sensors with LED component turning ON at regular intervals (peaks shown at around 50mA), indicating no fault in the LED component.

Thus, we can use current signatures to accurately detect and isolate faults. Furthermore, the signature remains same across sensors from the same manufacturer and is also agnostic to the environment in which it is deployed. Hence, a current signature collected once is sufficient to detect faulty sensors in the deployment.

3 IMPLEMENTATION AND DEMO

The typical operation for fault detection is as follows: Before deploying, we record the current signature of a working sensor and store it in the microcontroller EEPROM. When the devices are deployed, the current signatures of the $PM_{2.5}$ sensor under test are collected regularly. We then classify each signature as working or faulty by using template matching with the correlation coefficient as a measure of similarity between the reference fingerprint (template stored in EEPROM) and the collected fingerprint. If the similarity metric is above a user-defined threshold, then the sensor is said to be working or vice-versa.

In this demo, we highlight the issue of data faults arising from component failures in $PM_{2.5}$ sensors deployed in the real-world. We then introduce and showcase the use of current signatures to detect and isolate faults in $PM_{2.5}$ sensors. For demonstration purpose, we have fabricated a low-cost add-on custom board with a current sense amplifier for current measurement [3]. The add-on board measures and evaluates the current signature to determine the sensor status and can be easily retrofitted with $PM_{2.5}$ sensor as shown in Figure 4(a) and (b). Thus, policymakers can now use reliable data based on the derived sensor status for taking decisions.

REFERENCES

[1] 2020. Particulate Matter Sensor SPS30. https://www.sensirion.com/en/environmental-sensors/particulate-matter-sensors-pm25/.

[2] 2020. The Plantower PMS5003 and PMS7003 Air Quality Sensor. http://www.plantower.com/en/content/?110.html.

[3] 2020. Texas Instruments INA4180 Quad-channel, 26V current sense amplifier for cost-sensitive systems. http://www.ti.com/product/INA4180.

[4] Tusher Chakraborty, Akshay Uttama Nambi, and et al. Chandra, Ranveer. 2018. Fall-curve: A novel primitive for IoT Fault Detection and Isolation. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. ACM SenSys.

[5] Ni et al. 2009. Sensor network data fault types. *ACM Transactions on Sensor Networks (TOSN)* 5, 3 (2009), 25.

[6] WHO. 2020. 7 million premature deaths annually linked to air pollution. https://www.who.int/mediacentre/news/releases/2014/air-pollution/en/.