A Data Generation Approach for Intelligent Fault Detection and Diagnosis in District Heating

Jonne van Dreven[1,2], Veselka Boeva[1], Shahrooz Abghari[1], Håkan Grahn[1], Jad Al Koussa[2] and Emilia Motoasca[2]

I. INTRODUCTION

Around 46% of the total produced energy worldwide is used for heating and cooling purposes. Heat production is responsible for 40% of global Carbon Dioxide (CO2) emissions [1]. District Heating (DH) systems are crucial in sustainable energy usage and can play an important role in reducing greenhouse gas emissions and mitigating the effects of climate change. DH systems provide heat for residential and commercial buildings in many countries, such as Denmark (around 65%) and Sweden (above 45%) [2]. DH systems are complex as they involve a wide range of interconnected components, including heat exchangers, pumps, valves, and pipes and are prone to various faults that can negatively impact their performance, energy consumption, customer comfort, and maintenance costs.

Fault Detection and Diagnosis (FDD) is a critical task in DH systems to ensure that occurring faults have a minimal impact on the system operation. Traditional FDD methods rely on manual inspections, which can be time-consuming, expensive, and impractical. In recent years, Machine Learning (ML) and Artificial Intelligence (AI) approaches have emerged as promising tools to enhance FDD in DH systems. These methods can analyze large amounts of data and detect faults close to real-time, providing DH operators with timely and accurate information for maintenance and helping reduce the impact of faults.

One of the major challenges of intelligent FDD in the DH domain is the lack of labeled data on machine health state [3], as it is a complex, time consuming and expensive task. The lack of standardization and uniformity in data collection makes the process more difficult. Labeled data refers to data that has been classified with information about the health state of a system, and is crucial for the training of supervised ML models. This paper describes an approach for labeled data generation in DH, using DH emulations while inducing physical faults.

II. OBJECTIVES

The main objective of the study is to generate realistic (faulty) DH data using emulations that can be used for intelligent FDD. Emulation refers to the process of creating a physical replica of a system, and the typical functionality of the original system, rather than just simulating its behavior. Emulations seek to replicate a system exactly as it is in the real world without any simplifications or abstractions, i.e., emulations can be more accurate than simulations since they can reproduce the complexities and nuances of real-world DH systems.

The study aims to validate and enhance the experimental setup, while improving the data generation method.Furthermore, the study endeavors to investigate the feasibility of incorporating additional faults, as outlined in Table II, in order to construct a more comprehensive data set for intelligent FDD. The output of this study can be employed for a number of subsequent studies, e.g., the optimal data helps learn health profiles, which can be assessed for distinguishing normal and faulty behavior. If promising, these profiles can help in real-world data labeling or validation purposes.

III. PROBLEM STATEMENT

We presented a comprehensive literature review in [3], which concluded that one of the major challenges in DH is the lack of labeled data on machine health state.

A substation is a key component in DH system that acts as an interface to distribute heat to buildings. It connects the DH network (primary side) to the local distribution pipes (secondary side) that supply heat to individual buildings. Substations typically contain a range of components, including heat exchangers, pumps, valves, and control systems. The heat exchangers are used to transfer heat from the primary side to the secondary side, while pumps and valves are used to control the flow of the energy/heat medium (usually hot water) through the system. Control systems are used to monitor and manage the operation of the substations, ensuring that the heat is transferred efficiently and reliably to the connected buildings in accordance to their space heating (SH) and domestic hot water (DHW) needs, i.e., heat demand profiles.

A substation typically incorporates a heat meter which is designed to automatically monitor the heat consumption in order to facilitate accurate billing. This device is comprised of two temperature sensors, a flow meter, and a calculator. The temperature sensors measure the temperature of water in the primary supply and return pipes. The flow meter measures the flow at the primary side, before the point where SH and DHW hydraulic circuits split, and it is usually placed on the return pipe. The metering instruments can be powered

[1]Department of Computer Science, Blekinge Institute of Technology (BTH), Karlskrona, Sweden jonne.van.dreven@bth.se

[2]Unit Energy Technology, Flemish Institute for Technological Research (VITO), Genk, Belgium

from the electric power grid or batteries, depending on the specific implementation. The calculator is responsible for performing various calculations on the data, to derive some parameters such as energy. Table I summarises the typical variables a heat meter collects.

TABLE I
VARIABLES COLLECTED BY A TYPICAL HEAT METER.

Feature	**Notation**	**Unit**
Primary supply temperature	T_{ps}	°C
Primary return temperature	T_{pr}	°C
Volume flow	$\dot{V}$	m^3
Accumulated volume	V	m^3
Accumulated energy	Q	J

DH substations are vulnerable to various types of faults, including mechanical failures, electrical issues, and sensor malfunctions. Some studies suggest that between 43% and 75% of tested DH substations (in Sweden) run sub-optimally due to faults [4], [5]. Faults can cause a significant decrease in heating efficiency, and increase operational costs. Therefore, promptly detecting and diagnosing faults in DH substations can reduce their impact and ensure efficient operation. Table II presents the most common faults in substations.

TABLE II
COMMON FAULT OCCURRING IN DH SUBSTATIONS [6].

Number	**Fault types**	**Occurrences**
1	Leakages	33%
2	Customer internal system	31%
3	Control valve	13%
4	Actuators	10%
5	Control system and controller	5%
6	Inferior gaskets	5%
7	Heat exchanger	3%

IV. METHOD

Fig. 1 depicts the schematic of the substation designated for the emulations in a laboratory set-up. The study aims to generate realistic time series data with a five-minute interval for a typical winter week in Belgium with a constant supply temperature of $70°C$. The weather model represents a typical winter week in Belgium. Sensors are placed on the primary and secondary sides, collecting the same variables as seen in Table I. The set-up incorporates a simulated heat consumption model for a standard residential building together with a physical heat consumption model—a climate chamber. The main focus of the experiment is exploratory in nature, with two types of emulations planned: 1) optimal substation behavior and 2) faulty substation behavior. The faulty condition will constitute a stuck control valve (see Fig. 1). The selected fault is chosen as it is common [6] and deterministic to induce compared to more dynamic faults such as leakages or gasket failures. Our aim is to collect two data sets, one containing optimal and the other faulty behavior. Future experiments may incorporate additional faults from Table II.

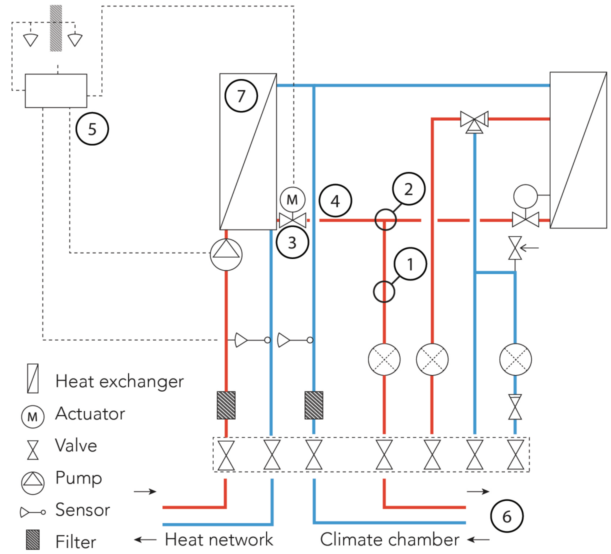

Fig. 1. Experimental setup of a substation at VITO. Numbers refer to faults from Table II.

V. CONCLUSION

In conclusion, district heating (DH) plays an essential role in sustainable energy usage, but faults have several negative impacts on the system, such as reduced performance or increased operational costs. Machine learning and artificial intelligence approaches have emerged as promising tools to enhance fault detection and diagnosis (FDD) in DH systems. However, one major challenge is the lack of labeled data on the machine health states. This paper proposes an approach for generating labeled data for a faulty DH substation using emulations while inducing physical faults in a controlled laboratory set-up. Our objective is to gather two data sets of a typical winter week in Belgium, to analyze in details whether the emulations are realistic and appropriate for intelligent FDD methods.

REFERENCES

[1] "Heat – Renewables 2019 – Analysis," IEA. https://www.iea.org/reports/renewables-2019/heat (accessed Feb. 28, 2023).

[2] "District Heating – Analysis," IEA. https://www.iea.org/reports/district-heating (accessed Mar. 02, 2023).

[3] J. van Dreven, V. Boeva, S. Abghari, H. Grahn, J. Al Koussa, and E. Motoasca, "Intelligent Approaches to Fault Detection and Diagnosis in District Heating: Current Trends, Challenges, and Opportunities," Electronics, vol. 12, no. 6, Art. no. 6, Jan. 2023.

[4] H. Gadd and S. Werner, "Fault detection in district heating substations," Applied Energy, vol. 157, pp. 51–59, Nov. 2015.

[5] S. Månsson, K. Davidsson, P. Lauenburg, and M. Thern, "Automated Statistical Methods for Fault Detection in District Heating Customer Installations," Energies, vol. 12, no. 1, Art. no. 1, Jan. 2019.

[6] S. Månsson, P.-O. Johansson Kallioniemi, M. Thern, T. Van Oevelen, and K. Sernhed, "Faults in district heating customer installations and ways to approach them: Experiences from Swedish utilities," Energy, vol. 180, pp. 163–174, Aug. 2019.