

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

Jonne van Dreven^{1,2}, Veselka Boeva¹, Shahrooz Abghari¹, Håkan Grahn¹, Jad Al Koussa² and Emilia Motoasca²

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

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

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

