DOI 10.20357/B7W02S

BERKELEY LAB

Lawrence Berkeley National Laboratory

What's in a Name? Developing a Standardized Taxonomy for HVAC System Faults

Yimin Chen, Eliot Crowe, Guanjing Lin, and Jessica Granderson

Lawrence Berkeley National Laboratory

Energy Technologies Area

August 2020

What's in a Name? Developing a Standardized Taxonomy for HVAC System Faults

Yimin Chen, Eliot Crowe, Guanjing Lin, and Jessica Granderson
Lawrence Berkeley National Laboratory

ABSTRACT

Faults occurring in heating, ventilation and air-conditioning (HVAC) systems have significantly negative impacts on building energy consumption, occupant comfort, and indoor air quality. In the past thirty years, extensive research has been conducted on fault detection and diagnostics (FDD) methods, and there are now dozens of commercially available FDD software tools. Growing adoption of FDD tools has the potential to generate a massive and useful data set on fault characteristics. However, the lack of a unifying taxonomy is a significant barrier to efficient analysis and evaluation of FDD outputs. Therefore, there is a strong need to develop a robust taxonomy which can better represent and interpret FDD output data.

This paper documents the development of a unifying taxonomy for HVAC system faults in commercial buildings, with initial focus on air handling units, variable air volume terminal units, and roof top unit systems. The developed fault taxonomy employs both a physical hierarchy of HVAC equipment and a cause-effect relationship model as tools to better understand and support root cause analysis for HVAC faults. A variable air volume terminal unit is used as an example to demonstrate the application of the developed fault taxonomy. The taxonomy has short-term application in a major U.S. study on fault prevalence, and promises longer term benefits to FDD software developers and building operators by creating a foundation for improved approaches to identifying and resolving HVAC faults.

Introduction

In the United States (U.S.), buildings consumed around 40% of primary energy and 70% of electricity consumption in 2018 (EIA 2019). Significant energy savings can be achieved by integrating state-of-art technologies such as optimal building control strategies, building data analytics, and smart maintenance (Shaikh, Nor et al. 2014). With rapid development of data collection and communication technologies, data generated from the entire lifecycle of a building such as design, installation, operation, and maintenance can be more efficiently collected. This data may be employed through data analytics software to improve building operation reliability, increase occupants' thermal comfort, reduce energy consumption, and enhance building's capabilities to interact with electric grids and renewable energy systems (Blum, Lin et al. 2018).

Fault Detection and Diagnostics (FDD) software is a building analytics tool that has seen significant growth and attention over the past decade. FDD tools analyze building operational data to determine whether a system operates abnormally, as well as to support root cause analysis efforts. These tools apply sophisticated algorithms to building automation system (BAS) data to produce detailed fault reports. For example, whereas a BAS may alert users to a single operational point being outside of a specified range, an FDD tool can analyze many points in combination to determine that the system as a whole is operating sub-optimally. Recent data from a United States government funded partnership (Smart Energy Analytics Campaign) with commercial building owners has demonstrated a median savings of 9% whole building energy savings from FDD deployment (Kramer et al. 2019) In the past two decades, extensive research efforts have been conducted on the development of advanced approaches to improve data analytics capabilities of FDD tools in buildings (Brown, Walter et al. 2014). The number of commercial FDD tools has been rapidly growing, with at least 32 tools available in the U.S. as of 2020 (Lawrence Berkeley National Laboratory 2020).

The rapid expansion in deployment of FDD software is generating and storing operational data at a scale orders of magnitude beyond what was occurring just a decade ago. This, in theory, offers an opportunity to understand the nature of commercial buildings' performance at an unprecedented level of granularity. However, there are several fundamental barriers to meta-analysis of FDD data, including:

- Inconsistent fault naming conventions;
- A mixture of fault type definitions with opaque relationships between them;
- Lack of a consistent physical hierarchy that can be used to classify faults occurring at different levels of operation (e.g., component level, sub-system level, whole system level).

Inconsistent data architecture and naming conventions will significantly reduce the value of FDD-generated data. Data labels, structure, class, type and other features are very important for a consistent data representation, which can be used to enable efficient operations and analytics on various datasets (Chen, Mao et al. 2014). Some attempts have been made to address this issue in the building controls industry. For example, *Brick* provides a uniform metadata schema which defines a concrete ontology for sensors, subsystems and the relationships between them (Balaji, Bhattacharya et al. 2018). Another example is Project Haystack (Charpenay, Käbisch et al. 2015) which uses standardized tags to label different entities, i.e., Site (single building with its own street address), Equip (physical or logical piece of equipment within a site), and Point (sensor, actuator or setpoint value for an equipment). In both attempts, semantic data models have been developed to dramatically increase the efficiency of data analytics from a diversity of data analysts such as building facility operators, maintenance providers, utility companies and so on.

While some efforts have been made to classify FDD outputs, none have been comprehensive to the degree necessary to enable effective meta-analysis of HVAC faults at national scale. With the increasing deployment of FDD tools in the marketplace, a high volume of fault data is being generated across equipment/systems in different types of buildings. These FDD tools exhibit very different fault naming and indexing conventions for the faults they identify. In addition, various fault definitions may be used across FDD software tools (Frank, Lin et al. 2019). Further, FDD reports lack vital information and standardized name labels.

In this paper we describe the development of an HVAC fault taxonomy that formalizes fault naming conventions, resolves issues arising from the mixing of different fault type definitions, and presents a consistent physical hierarchy within which fault relationships can be better understood. The taxonomy was principally designed to support a large-scale U.S. fault prevalence study, but will have broader application for HVAC researchers, FDD software developers, and commercial building owners. This paper is organized as follows:

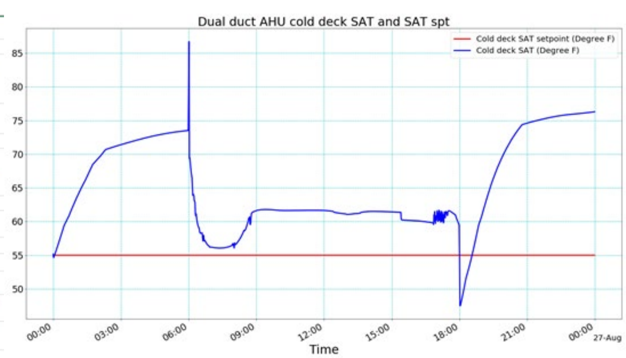
- Background section covers the state of the art in HVAC fault classification, FDD software development, and barriers to efficient analysis of FDD tool outputs;
- Taxonomy Development Methodology will describe the key elements of the HVAC fault taxonomy and how each element was developed to address unique considerations;
- Example Application of Taxonomy provides a practical example of how the taxonomy is applied to a variable air volume (VAV) terminal unit;
- Discussion will cover the implications of the fault taxonomy and related challenges;
- Conclusions and Future Work will summarize the key outputs of the taxonomy development effort, short-term applications, its broader long term significance.

Background

FDD Software Description

Existing FDD software tools employ HVAC system trend data from BAS (sometimes in combination with meter data, weather data or other information) to implement FDD algorithms (e.g., BAS instructs AHU outside air damper to be at minimum position, but FDD algorithm identifies that mixed air temperature tracks the outside air temperature, indicating that the outdoor air damper is actually stuck in an open position) and support HVAC system operating management (Granderson, Lin et al., 2018). FDD tools can output their analytics results in a variety of formats, including tabular/text and visual format, as shown in Figure 1. In the tabular/text summary example (Figure 1a), a semantic description of a fault is given. A semantic representation of systems/equipment/components operational status is used to describe a fault. Each description can be viewed as a fault message log and be recorded with different time duration in the database. To support root cause analysis of reported faults, FDD tools may also offer charting and trending capabilities to users (Figure 1b). The underlying fault detection algorithms may be the same in either case; these examples relate specifically to the presentation of FDD results. Real case study

Building	Equipment	Component	Analysis	Start Date	Notes Summary
Building #1	HVAC Boiler	Boiler1	Heating plant	8/22/2018	Boiler efficiency is low.
Building #1	HVAC AHU	AHU2	AHU coil	8/22/2019	Simultaneous heating and cooling. Return RH higher than setpoint. Leaking heating valve. Supply temp reset error.
Building #1	HVAC AHU	AHU2	AHU coil	8/22/2019	Occupancy not compliant with schedule.
Building #3	HVAC AHU	AHU4	AHU fan	8/22/2019	No static pressure reset.
Building #1	HVAC AHU	AHU2	AHU coil	8/22/2019	Simultaneous heating and cooling. Leaking cooling valve. Out of range sensor error (low). Supply air temperature short cycling.
Building #1	HVAC AHU	AHU1	AHU coil	8/22/2019	Excess mechanical cooling.
Building #1	HVAC AHU	AHU1	AHU coil	8/22/2019	Minimal load across loop. Low loop temp difference.
Building #2	HVAC AHU	AHU1	AHU economizer	8/22/2019	Economizer should be off.
Building #3	HVAC AHU	AHU4	AHU coil	8/22/2019	Simultaneous heating and cooling.
Building #3	HVAC AHU	AHU4	AHU control	8/22/2019	No static pressure reset.
Building #3	HVAC AHU	AHU4	AHU control	8/22/2019	Low loop temp difference. Supply/outlet temp mismatch. Configuration flag.
Building #1	HVAC AHU	AHU2	AHU sensor	8/22/2019	Heating on, economizer should be off. Out of range sensor error (low).



(a) (b)

Figure 1. Examples of FDD tool analytics outputs: (a) tabular/text summary (b) visual display showing time-series trend of supply air temperature (SAT) and SAT setpoint

Strictly speaking, FDD software does not truly "diagnose" the root causes of faults, e.g., the software may identify an inoperable damper but would not know exactly why the damper is inoperable. However, FDD software has proven very effective for flagging inefficient operation and supporting operator's efforts to hone in on the specific components or control-related issues that need to be addressed.

Variety of Fault Type Definitions used by Existing FDD Software

The definition of the term "fault" is key to understanding the analysis output of FDD tools. Existing literature and commercial FDD tools use three general categories of fault type definition based upon how the faults are presented: condition-based, behavior-based, or outcome-based (Frank, Lin et al. 2019). The categories and corresponding definitions are provided in Table 1. As an introductory example shown in Table 1, consider an air handling unit (AHU) with its cooling coil valve stuck open. The unit's faulted state may be defined by the unit's condition (the cooling coil valve is stuck open), behavior (the supply air temperature can't meet its setpoint), or outcome (the unit's chilled water consumption is more than expected). If, however, the unit is experiencing a call for cooling, it would still be considered faulted under the condition-based definition (the valve is still stuck), but not under the behavior-based definition (The extra cooling due to the cooling valve stuck is compensated by the heating provided by the heating coil at the same time, the supply air temperature meets setpoint). The unit's state under the outcome-based definition would be determined by the amount of chilled water flow through the stuck valve compared to an expected level of chilled water consumption. While condition-based and behavior-based fault conventions are more often seen in the tabular/text analysis summary of the commercial FDD tools outputs, all three definitions may be used in the tools depending on the available input data and the tool developer's preferred approach.

Table 1. Fault definition conventions

Category	Condition-Based Fault	Behavior-Based Fault	Outcome-Based Fault
Definition	Presence of an improper or undesired physical *condition* in a system or piece of equipment	Presence of improper or undesired *behavior* during the operation of a system or piece of equipment	A state in which a quantifiable *outcome* or performance metric for a system or piece of equipment deviates from a correct or reference outcome, termed the expected outcome
Example	The cooling coil valve is stuck open	The supply air temperature can't meet its setpoint	The chilled water consumption is more than expected

Barriers to Meta-Analysis of FDD Tool Analytics Outputs

As mentioned in the introduction, there are three fundamental barriers to meta-analysis of FDD tool analytics outputs, which drives a need to develop a unified HVAC fault taxonomy. Each of the three barriers are addressed in detail below:

Inconsistent fault naming conventions: Various fault names, which may refer to the same fault in the system, are used in different FDD tools. This issue may even arise with different versions of the same software tool, or with unique installations of any given tool. This causes an obvious discrepancy among different FDD reports or software. For example, in one commercial FDD tool, a "*discharge air damper hunting*" fault is reported to reflect a malfunctioning damper control in variable air volume (VAV) terminal units. But in another FDD tool, this fault may be reported as a "*discharge air damper cycling*" fault or an "*unstable damper*" fault. This inconsistency in naming does not inhibit the FDD software's effectiveness for an individual building (i.e., building operators will become familiar with the fault and the appropriate actions to take in response), but it is a significant barrier when wanting to gather a large set of meta-data for the purpose of characterizing faults at a national scale.

Mixture of fault definition conventions: Both condition-based and behavior-based fault conventions are widely used across FDD tools. However, the relationships between these two conventions are not sufficiently defined to identify causal relationships between those faults. Therefore, when using FDD tools' default naming conventions, redundant information may be used repeatedly. This may cause an over-count of faults and increasing difficulty to identify root causes in a complex system without an accurate model to unify different fault conventions. For example, a "*low supply air temperature*" fault, a "*simultaneous heating and cooling*" fault and a "*stuck cooling coil valve*" fault may be reported concurrently. At face value this would be counted as three faults, when in reality the stuck cooling valve may be the single true fault that is resulting in the other two effects being observed. A successful meta-analysis of FDD data requires the relationships between different fault types to be defined.

Lack of a consistent hierarchy that can be used to resolve faults occurring at different levels of operation: A study of faults occurring in commercial buildings may be viewed on many levels, for example by system type, sub-system, equipment, or individual components. However, there is no single hierarchy employed by FDD tools for the purpose of consistently identifying a fault's position within the system as a whole. Some key information is sometimes missing in the fault name generated by the FDD report. For example, a "*flat sensor*" fault may be reported by a FDD tool but no information on what type of sensor (e.g., temperature or air flow) is given in the fault name. Given that FDD tools include fault rules built upon data from different hierarchical levels, the lack of a formal hierarchy presents challenges when collecting and analyzing data from multiple tools and many buildings.

These three barriers significantly hinder effective large scale meta-analysis of FDD results. Addressing these issues was critical to a planned study of nationwide HVAC fault prevalence. Beyond that specific short-term need, standardized taxonomies have generally proven useful in a variety of contexts. For example, a taxonomy that resolves the above-mentioned barriers may support development of more optimized FDD software algorithms and standardized approaches to root cause analysis of faults. With those issues in mind, we propose a novel taxonomy to better represent HVAC system faults and hierarchy.

Development of Taxonomy

To date, the taxonomy covers three major system types: packaged rooftop units (RTUs) air-handling units (AHUs), and variable air volume (VAV) terminal units. These systems were chosen as collectively they cover the majority of U.S. commercial buildings' HVAC systems. Approximately 90% of buildings and nearly half of the commercial floor space is served by RTUs. In addition, AHUs are a key element of the built-up systems that are common in large commercial facilities. Further, the taxonomy development was targeted at the highest impact faults that are commonly encountered in these system types.

In order to address the identified needs and barriers summarized in the Introduction our fault taxonomy comprises four key elements:

1. Fault type definitions;
2. Equipment physical hierarchy;
3. Fault relation models;
4. Fault identification code scheme.

Each of these key elements is described in the following sections.

Fault Definition Conventions

FDD results in literature (Breuker and Braun 1998, Schein and Bushby 2006, Wang, Chen et al. 2012) and commercial FDD tools were reviewed and classified using the fault definitions presented in Table 1. For example, under such definitions, a "*reheat coil valve stuck*" fault is categorized as a condition-based fault, whereas a "*discharge air temperature abnormal*" fault is categorized as a behavior-based fault. This therefore helps identify the fault instances that may in some cases refer to a single fault state as opposed to multiple independent fault states and also makes it possible to develop a relation model to disclose the causal relations for the faults using different types of definitions. Condition-based faults and behavior-based faults are the focus of this study, since the use of outcome-based faults is relatively rare.

Physical Hierarchy and Fault Library

A complete HVAC equipment fault definition should contain information which can be used to locate the fault and reflect the fault feature. In order to reach this goal, we first define a system physical hierarchy, and then assign a unique fault name from a summary fault library.

An accurate physical hierarchy is critical to reflect various levels in a complex HVAC system. Some studies have been conducted to use various structures to represent the physical entities in a building. For example, Brick presents an integrated, cross-vendor representation of the multitude of subsystems in modern buildings which include HVAC, lighting, fire, security, and other systems (Balaji, Bhattacharya et al. 2016). The names we selected for the equipment and components are drawn from the Brick vocabulary library. However, we also defined some new names which Brick standard does not include. In the taxonomy development, four levels were adopted to locate a fault, as shown in Figure 2. The highest level defines the system type, for example, "HVAC system" (referred to as S1) is used at this level (for our study this is the only defined system. The next level identifies the specific equipment within the system. For example, AHU, RTU, VAV terminal unit etc., are grouped into this level (for our study these are the only equipment types in scope, but others could be added at a later date). Below the equipment level, specific components are defined under two separate levels, so that fault location can be accurately determined. Two levels are used to identify the component because it can differentiate the generic component type from the specific component type. For example, in the taxonomy a generic component type "sensor" (Component Level 1, referred to as C1) is used to cover all sensors used in a VAV terminal unit. Specific sensor types such as "discharge air temperature sensor" may then be defined under Component Level 2 (referred to as "C2"). Within this structure, "control sequence" is considered a Level 1 Component.

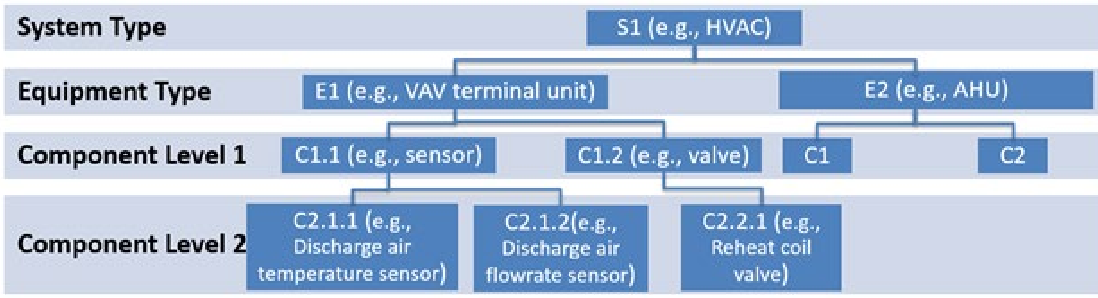


Figure 2. Physical hierarchy levels used in the taxonomy, with examples for each level

To address the challenge of inconsistent fault naming conventions, in this study, standard fault names are assigned to each component. For example, in the condition-based fault category for the component "damper," three fault names are defined in the taxonomy, i.e., "*damper leakage,*" "*damper stuck*", and "*damper unspecified*". Each of these three fault names may capture multiple faults as reported by FDD tools; for example, "*damper leakage*" represents fault semantic descriptions such as "*damper cannot be fully closed,*" "*damper is not sealing*", "*damper is not fully closed.*"

For behavior-based faults in the taxonomy, the fault name may represent either the abnormality in one physical variable or that a rule among different physical variables is broken. For example:

- "Discharge air temperature abnormal" (single physical variable);
- "Rule between outside air temperature and mixed air temperature is broken" (multiple physical variables).

Once fault naming conventions were established in the taxonomy, a fault library was developed for each of the three main system types (e.g., 19 condition-based faults and 5 behavior-based faults are defined for VAV terminal units).

Fault Relation Model

In FDD reporting, both condition-based faults and behavior-based faults can be found in the software outputs. For each condition-based fault, a set of related behavioral symptoms (behavior-based faults) could occur, and these relationships can be defined using a relation model. In this study, we developed a comprehensive relation model to connect the condition-based faults and associated behavior-based faults based on expert knowledge and in consultation with a project Technical Advisory Group.

It should be noted that a condition-based fault may propagate and affect different equipment or even subsystems due to the highly coupled HVAC system (e.g., a chilled water supply pump fault in the primary subsystem can cause multiple symptoms in downstream equipment such as an abnormal cooling coil valve position and abnormal supply air temperature in AHUs). Therefore, a condition-based fault and associated behavior-based faults may cross between different pieces of equipment or subsystems. However, in this study, only 'local' fault relation models with a single piece of equipment were developed.

Fault Identification Code Structure

A unique fault identification (ID) scheme was designed to label each fault, so that each fault in the library can be efficiently retrieved and used in different domains. For the fault ID, a unified tagging format is used to differentiate the fault by different character sections as described in Figure 3. The first section identifies the equipment name. Both the second section and the third section define the component name (one may be sufficient in some cases). This increases the physical granularity, and can help diagnose the root cause of a fault more efficiently. The last section identifies the fault. This fault ID is used to tag a fault so that it can exactly represent the fault type and location in a fault library. For example, the fault ID of fault “VAV terminal unit reheat valve stuck” is VAVUNIT-RHC-Vlv-Stuck, as shown in Figure 3.

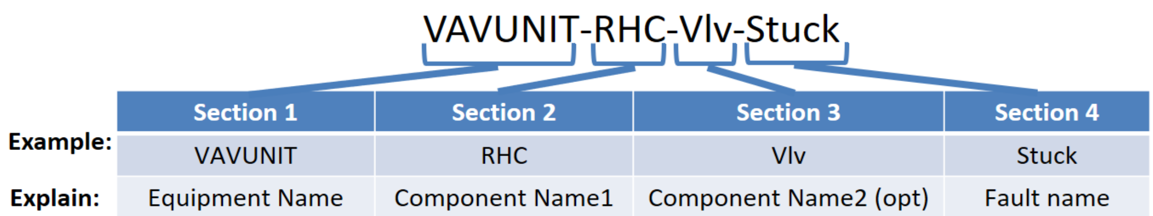


Figure 3. Example of tagging a condition-based fault identification code

Example: Taxonomy Application for VAV Terminal Unit Faults

In this section, we illustrate an example application of the HVAC fault taxonomy for VAV terminal unit faults, employing a typical VAV terminal unit configuration.

Physical Configuration of the VAV Terminal Unit

A VAV terminal unit (also known as a “VAV box”) is used to adjust the airflow rate and discharge air temperature dispatched into a conditioned space from an AHU. The control of a VAV terminal unit is based on the heating/cooling loads and ventilation requirement of the space served by the VAV terminal unit (Liu, Wen et al. 2014).

VAV terminal units can be configured in many ways to meet conditioning requirements. We employed a commonly-used physical configuration to develop the fault taxonomy (see Figure 4). In this configuration, the VAV terminal unit consists of a discharge air damper, a hot water reheat coil and coil valve, and two sensors (discharge air flowrate sensor and discharge air temperature sensor). The discharge air damper position is adjusted to introduce variable air flow in the zone, so that the zone temperature can meet the temperature setpoint. It is noted that the taxonomy will be extended under different equipment configurations because more components may be integrated.

Taxonomy of Condition-Based Faults

VAV terminal unit condition-based faults can be physically located using each level of the physical hierarchy. A graphical representation of the physical hierarchy of faults for VAV terminal units is provided in Figure 5. The hierarchy defines the types of components (sensor, damper, coil, etc.) at Component Level 1. Component Level 2 provides detailed component names, and the fault type applicable to each component category is shown on the bottom level of the hierarchy.

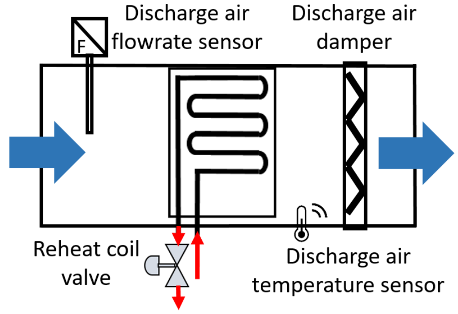


Figure 4. Typical VAV terminal unit configuration, as employed in the HVAC fault taxonomy

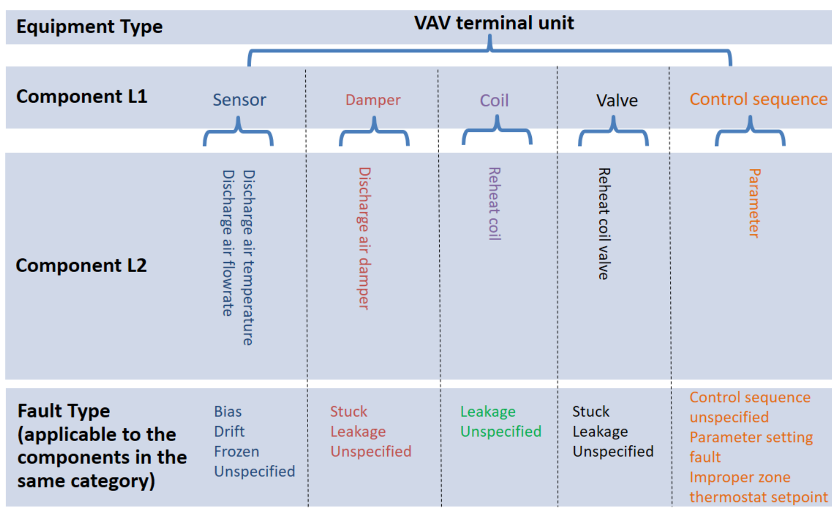


Figure 5. Physical hierarchy for VAV terminal faults (condition-based faults)

Table 2 provides an example of the developed taxonomy for a condition-based fault affecting a VAV unit reheat coil. In the table, the two-level component and corresponding component names are provided in the first two columns. A fault type term and corresponding description are given in the third and fourth column. In the fifth column, examples of other descriptions found in FDD tools are provided. A fault ID which can be used to transparently codify the fault data is given in the last column.

Taxonomy of Behavior-Based Faults

Figure 6 gives the hierarchy for the taxonomy for behavior-based faults applicable to the VAV terminal unit example. The physical variable types are given at Component Level 1. At this level, four measurements are provided: temperature, damper control, coil valve control and air flow. Various measurements are given at Component Level 2.

Table 2 Example of condition-based fault types, name and ID

Component Type (L1)	Component Type (L2)	Fault Type	Fault Name	Other Fault Descriptions	Fault ID
Coil	Reheat coil valve	Leakage	Reheat coil valve leakage	Coil valve is leaking by; coil valve has leakage; coil valve is leaking.	VAVUNIT-RHC-Vlv-Leak

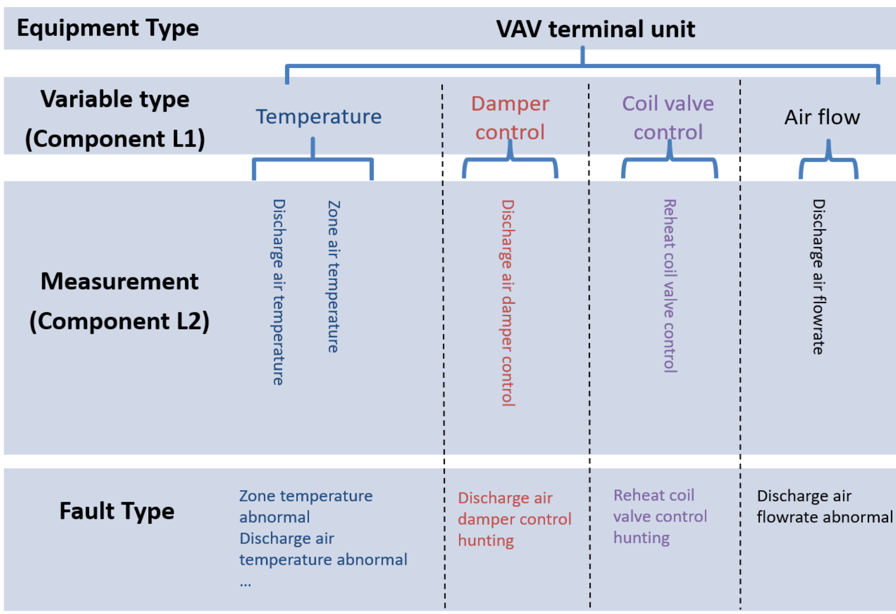


Figure 6. Physical hierarchy for VAV terminal faults (behavior-based faults)

Table 3 provides an example of the developed taxonomy for the behavior-based faults.

Table 3 Example of behavior-based fault types, name and ID

Component Type (L1)	Component Type (L2)	Fault Type	Fault Name	Other Fault Descriptions	Fault ID
Temperature	Discharge air temperature	Abnormal	Discharge air temperature abnormal	Discharge air temperature too high; discharge air temperature too low; discharge air oscillation, discharge air temperature swing.	VAVUNIT-DAT-Abnormal

Relation Model

The relation model can be illustrated with a tree diagram, with a simplified excerpt shown in Figure 7. Figure 7 indicates that in a VAV terminal unit, when a "*discharge air damper stuck*" condition-based fault occurs, it may cause multiple associated behavior-based faults such as "*discharge air flowrate abnormal,*" and "*zone temperature abnormal.*"

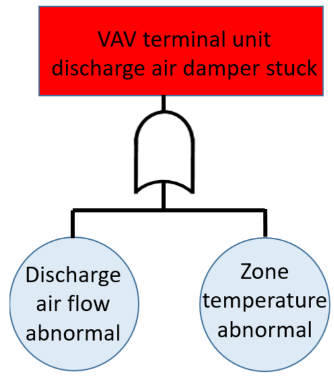


Figure 7. Illustration of a relation model

Through the application of the taxonomy, data from various FDD tools may be unified into a structure such that faults are defined both physically and in causal relation to each other, employing a consistent and transparent naming scheme that allows for effective meta-data analysis.

Example Taxonomy Application for a Sample FDD Tool Report

To assess the applicability of the developed taxonomy we mapped faults reported by a commercial FDD tool. We obtained an example fault report in a tabular format covering five days from five buildings across the U.S. From the semantic text fault list generated from the report, we identified a total of seventy-one fault names, which included 36 condition-based fault names, 34 behavior-based fault names, and 1outcome-based fault name for the buildings' AHUs. Using the fault taxonomy, 33 condition-based fault names, 28 behavior-based fault names and 1 outcome-based fault name in the FDD tool's diagnostic report can be successfully mapped (87% of the reported faults). The reason of unsuccessfully mapped names will be discussed in the next section. Table 4 lists the mapping results.

Table 4. Fault name mapping results

	Total	Condition-Based	Behavior-Based	Outcome-Based
Total fault names from FDD tool report	71	36	34	1
Number of faults successfully mapped to taxonomy	62	33	28	1
Successful mapping percentage	87%	92%	82%	100%
Consolidated number of unique fault names based on taxonomy mapping	35	16	18	1

As shown in Table 4, a total of 62 faults from the FDD sample report were mapped to 35 unique fault names in the taxonomy. This consolidation is a result of cases where multiple faults were mapped to a single fault in the taxonomy.

Discussion

Unifying fault taxonomy libraries for three common system types including AHU, RTU and VAV terminal units were developed based on careful analysis of academic literature and several example reports from FDD tools. The developed taxonomy is being deployed in support of a large-scale empirical study on the prevalence of faults in US commercial buildings. That activity will further verify the taxonomy's completeness and applicability using a large dataset sourced from several commercially available FDD tools. Based on the fault prevalence study it is expected that certain refinements will be made to the taxonomy, prior to its public release. Anticipated challenges and refinement options are discussed below.

Extension and Evolution of Fault Taxonomy

When developing the unifying fault taxonomy, the researchers included the most common faults reported in existing academic literature, and also reviewed the range of typical/common faults reported by commercial FDD tools. The selected equipment types (AHU, RTU and VAV terminal units) are widely deployed in today's commercial buildings' HVAC systems. However, different system physical configurations and additional components may be adopted in each type of equipment to meet different application requirements. For example, some AHUs are equipped with additional components such as heat recovery wheels and carbon dioxide sensors to improve energy efficiency and indoor air quality. Those components (and associated faults) are not included in the existing physical hierarchy. The fault prevalence study and other example applications will provide an opportunity to quantify the taxonomy's coverage, i.e., what portion of reported faults can be translated into the taxonomy structure? For those faults that cannot be translated into the taxonomy, decisions on updating the taxonomy can be made based on the frequency of non-covered faults, energy impact, the level of complexity that would be required in updating the taxonomy, etc.

Practical Applicability of Relation Model

The relation models in the fault taxonomy, which reflect the cause and effects relationships between condition-based faults and behavior-based faults, were developed to represent relationships within one piece of equipment of the system. Application of the relation model on large datasets from multiple data sources will provide valuable insights on its effectiveness in several dimensions:

- Will the relation model effectively handle the mixture of condition-based and behavior-based faults reported across a population of buildings, in a way that supports improved root cause analysis?
- What is the optimal way to quantify fault occurrence/prevalence for a population of buildings, given the inter-relationship between faults at different component/equipment levels and the mixture of condition-based and behavior-based faults?

- In what ways can the relation model be applied to a fault dataset to produce helpful insights for FDD software developers and building operators?

Further, HVAC systems are highly interactive, and a fault in one piece of equipment may affect other equipment in the system. For example, an AHU outdoor air damper stuck at too high a position in the summer season may cause pump speed to be abnormally high in the chiller plant. For future evolution of the taxonomy it could be beneficial for the relation models to be updated to include cross-equipment faults, so that the root-cause diagnosis of behavior-based faults can be more comprehensively covered. More comprehensive root cause analysis and risk assessment techniques may also be helpful in improving the relation model or providing supplementary resources to building operators.

Limitation of Naming and Tagging Scheme

The fault ID numbering scheme was developed after comparing naming approaches in existing fault reports and other naming conventions such as the Project Haystack system. These semantic data models make it easier to map complex FDD results and descriptions into a concise representation. However, we found that for some behavior-based faults, the naming and tagging scheme was vague. For example, when naming a fault where the rule between multiple measurements was broken, the taxonomy allows for only two measurements in the naming format. This may cause conflicts when more measurements are included. For example, we use "*mixed air temperature abnormal*" fault name to represent "*mixed air temperature should be between outside air temperature and return air temperature*" fault name which actually includes three physical measurements. Test applications of the taxonomy will help to fully uncover these kinds of gaps (and their impact), to support future taxonomy development efforts.

An overarching consideration for the taxonomy is long-term governance. Several areas of future taxonomy evolution are discussed here, but buildings (and their associated systems) are ever-evolving. Within the scope of the U.S. DOE-funded fault prevalence study the taxonomy will evolve to a certain degree, but beyond that there may be benefit in establishing an ongoing stakeholder-driven effort to maintain and manage future evolution of the taxonomy and related resources.

Conclusions and Future Research

The successful applications of data-driven or big data techniques rely on how data is represented and whether the data can be efficiently unified and interpreted. In this paper, we present a unifying taxonomy for HVAC system faults affecting AHUs, RTUs, and VAV terminal units. Built upon four key elements, the developed taxonomy gives an accurate and orderly classification of HVAC equipment faults based upon their characteristics and causal relations. This taxonomy is a foundational enabler for a major U.S. study on fault prevalence, and is expected to offer benefits to FDD software developers and operators by standardizing fault naming, system physical hierarchies, and bringing greater transparency to the relationships between faults.

Future work (beyond completion of the fault prevalence study) will include extension of the fault taxonomy library so that more equipment, faults, and corresponding physical configurations can be included. Further, more comprehensive analysis of cause and effect relationships and relative levels of risk could improve the taxonomy's utility for supporting ongoing root cause diagnostics. The nature of commercial buildings' HVAC faults has been studied for decades; the provision of a unified, transparent taxonomy will accelerate those efforts through common terminology and a more formal definition of physical and causal relationships.

Acknowledgement

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Office, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., et al. (2016)."Brick: Towards a unified metadata schema for buildings". Paper presented at the Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments.

Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., et al. (2018). "Brick: Metadata schema for portable smart building applications". Applied energy, 226, 1273-1292.

Blum, D., Lin, G., Spears, M., Page, J., et al. (2018). "When data analytics meet site operation: Benefits and challenges": eScholarship, University of California.

Breuker, M. S., & Braun, J. E. (1998). "Common faults and their impacts for rooftop air conditioners". Hvac&R Research, 4(3), 303-318.

Brown, R. E., Walter, T., Dunn, L. N., Custodio, C. Y., et al. (2014)."Getting real with energy data: Using the buildings performance database to support data-driven analyses and decision-making". Paper presented at the Proceedings of the ACEEE summer study on energy efficiency in buildings.

Charpenay, V., Käbisch, S., Anicic, D., & Kosch, H. (2015)."An ontology design pattern for iot device tagging systems". Paper presented at the 2015 5th International Conference on the Internet of Things (IOT).

Chen, M., Mao, S., & Liu, Y. (2014). "Big data: A survey". Mobile Networks and Applications, 19(2), 171-209. doi:10.1007/s11036-013-0489-0

EIA. (2019). Monthly energy review april 2019. Retrieved from https://www.eia.gov/totalenergy/data/monthly/previous.php:

Frank, S., Lin, G., Jin, X., Singla, R., et al. (2019). "A performance evaluation framework for building fault detection and diagnosis algorithms". Energy and Buildings, 192, 84-92.

Granderson, J., Lin, G., Singla, R., Mayhorn, E., et al. (2018). "Commercial fault detection and diagnostics tools: What they offer, how they differ, and what's still needed".

Kramer, H., Lin, G., Curtin, C., Crowe, E., et al. (2019). "Building analytics and monitoring-based commissioning: Industry practice, costs, and savings". Energy Efficiency, 1-13.

LBNL. (2020). Lawrence Berkeley National Laboratory. Retrieved from https://smart-energy analytics.org/product-service

Liu, R., Wen, J., Zhou, X., Klaassen, C., et al. (2014). "Stability and accuracy of variable air volume box control at low flows. Part 2: Controller test, system test, and field test". 20(1), 19-35.

Schein, J., & Bushby, S. T. (2006). "A hierarchical rule-based fault detection and diagnostic method for hvac systems". Hvac&R Research, 12(1), 111-125.

Shaikh, P. H., Nor, N. B. M., Nallagownden, P., Elamvazuthi, I., et al. (2014). "A review on optimized control systems for building energy and comfort management of smart sustainable buildings". Renewable and Sustainable Energy Reviews, 34, 409-429.

Wang, H., Chen, Y., Chan, C. W., & Qin, J. (2012). "An online fault diagnosis tool of VAV terminals for building management and control systems". Automation in Construction, 22, 203-211.