SYSTEMS-THEORETIC PROCESS ANALYSIS (STPA) IN BUILDING ENERGY RISK MANAGEMENT

Stylianos K. Karatzas and Athanasios P. Chassiakos
University of Patras, Patras, Greece

Abstract

As building energy management becomes a complicated process, traditional risk analysis is not adequate to address the management needs and advanced complex modelling techniques are required to handle the multi-dimensional synthesis of risks in the electricity systems. STPA (System-Theoretic Process Analysis) is a new hazard analysis technique using concepts of system and control theory. In this paper, the development of a risk management structure in the field of building energy management is presented. The hazard analysis process based on SafetyHAT information tool is applied for building operation 'accident' prevention, by illustrating the usage of the proposed tool. For this purpose, a case study considering a holistic energy management system in the tertiary building sector is examined.

Introduction

Risk management in the energy sector tends to become indispensable in energy applications nowadays. The inclusion of additional variables into the energy ecosystem makes it imperative to consider a reliable risk management framework. The aim of this study is to develop a general risk management methodology to the energy environment as a holistic approach to tackle risks associated with building energy management.

Within the scope of the study, potential "accidents" are analysed in a building energy ecosystem. An accident is defined as an undesired and unplanned event that results in a loss (including loss of human life or injury, property damage, environmental pollution, and so on) (Leveson, 2004). There are several available research methods in the bibliography, each one being distinguished by its own features and tools and applied in different fields. A typical model classification, considering also their evolution in time, results into three major groups, namely Linear or Sequential Models, Epidemiological Models and Systemic Models (Wienen et al. 2017).

On the way to define a fully-fledged methodology for modelling risks & hazards in a dynamic system, research shows that Traditional Hazard Analysis Methods work well for losses caused by failures in simple systems but are limited in their capability to explain accident causation in the more complex systems. They cannot handle with:

- Component interaction accidents
- Systemic factors (affecting all components and barriers)
- Software and software requirements errors
- System design errors
- Indirect or non-linear interactions and complexity
- The application of risk management methodologies in the building energy sector is limited and incomplete.

Towards an approach for complex systems safe operation, system theory basics are defined. According to system theory the system is treated as a whole, not as the sum of its parts. Also, relations and interactions among system components are considered and a primary concern is emergent properties, which are properties that are not in the summation of the individual components but "emerge" when the components interact, considering overall safety as an emergent property. In this direction , the relatively new Leveson's Systems-Theoretic Accident Model and Processes (STAMP) and Systems Theoretic Process Analysis (STPA) model, endeavor to model the dynamics of complex sociotechnical systems.

From STAMP to STPA

Systems-Theoretic Accident Model and Processes

Systems-Theoretic Accident Model and Processes (STAMP) is a relatively new methodological risk management framework which is considered in this study. The STAMP model (Leveson, 2002, 2004 & 2011) gives emphasis to the security restrictions and considers an accident in a complex system not just as the case of failure of some individual system component but rather the result of either an external factor or a malfunction within the system which has not been effectively addressed by the control system (Leveson, 2011). The model differentiates from the traditional approach in considering an accident as a sequence of events and as the result of insufficient control and ineffective application of constraints on the development, design and operation of the system (Ouyang et al., 2010). The new accident model is based on system and control theory rather than

reliability theory and safety is viewed as a control problem rather than a component reliability problem. A hierarchical safety control structure is used in STAMP to represent the system and control loops in it, showing how constrains are enforced. A typical control loop is presented in Figure 1.

Instead of addressing accidents as the results of event-chain, they are considered to result from a lack of constrains on behaviour at each level of a socio-technical system. The initial system design needs to impose appropriate behavioural constraints to ensure safe operation.

Systems Theoretic Process Analysis

System-Theoretic Process Analysis (STPA) is a hazard analysis method based on the STAMP accident model. STPA is based on system control theory and not on reliability theory established in most existing risk analysis techniques. The basic principles and characteristics of STPA are (Leveson, 2013; Friedberg et al., 2017; Horney, 2017):

- The best way to detect accident chances in complex systems is to omit causal factors that are not stochastic or for which no information is available. Probabilistic analysis results may not accurately reflect the actual risks and can be riskily misleading.
- Unlike traditional risk analysis techniques, STPA is stronger in identifying risk causes and hazardous scenarios, especially those related to system design and human behaviour.
- Because STPA is a top-down approach, system security engineering can be used early in the system development process to create high-level security requirements and constraints. If risk analysis is applied promptly and in accordance with planning decisions, the number and the cost of operating failures become negligible.
- STPA, supporting hierarchical safety control structures, can be used for both technical design and organizational planning.

STPA process can be divided into four phases with interconnected activities and can be considered as a repetitive process constantly backed up as the system design evolves. The individual procedures are:

1. Determination of the base system (system foundation).
2. Identification of potential Unsafe Control Actions (UCAs).
3. Development of constraints and limitations based on UCAs.
4. Determination of Causal Factors that may lead to UCAs.

STPA can be applied to any emergent system property in the system engineering and product lifecycle, not just safety. Also STPA system safety analysis can be integrated into the entire system engineering process resulting in a significant decrease in the cost of engineering for safety as well as in effectiveness and fewer losses. It can also reduce rework, which reduces cost and schedule. Figure 2 shows a simplified version of the standard system engineering V-model. This figure is used to illustrate how to integrate STPA into the standard system engineering process. The potential roles for STPA are shown in red. STPA can be used throughout the standard system engineering process, starting in the earliest concept development stage and contribute to all the activities in system engineering.

Information tool for the STPA methodology

In this section, an information tool is presented which models and records the STPA activities in a standardized way. There are various software tools that support risk analysis based on the STPA model, e.g., A - STPA (Krauss et al., 2015), XSTAMPP (Abdulkhaleq & Wagner, 2015) and SAHRA (www.sahra.ch).

In the present study, the SafetyHAT modelling tool, developed by the US National Transportation Systems Center (SafetyHAT User Guide, Becker & Van Eikema Hommes, 2014) is applied due to its simplicity and maturity in modelling and mapping the risk management diagrams as defined in the STPA methodology. The basic features for the SafetyHAT tool are the following:

- The SafetyHAT directs analysts based on preparatory and analytical steps of STPA to provide an improved data input process, standardize information input to the system and adapt to multiple application fields.
- The SafetyHAT exploits the dynamics of a relational database for organizing and managing large amounts of data. It can support storage of big analysis data (a control system can generate more than 10,000 entries) and enhance data integrity while modifying or deleting data.
- SafetyHAT can facilitate the documentation of risk analysis through detailed presentation of STPA analysis results, documented final results and data sharing and reuse.
- System analysis capability can be extended based on previous models and analysis results.
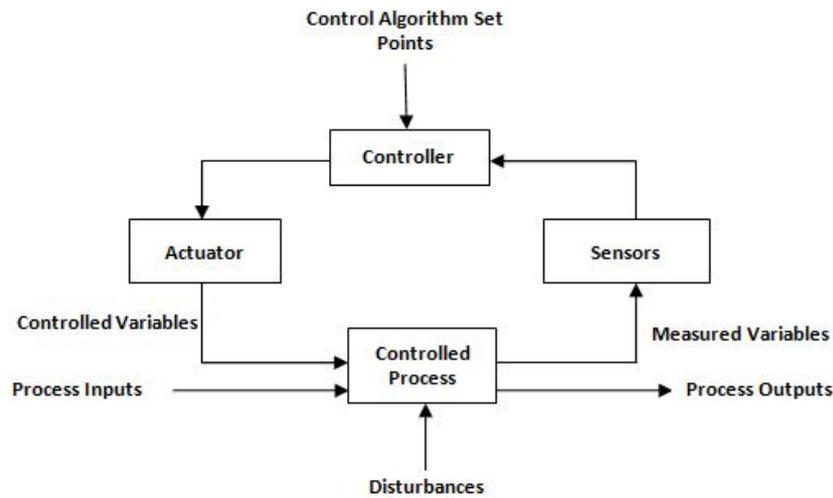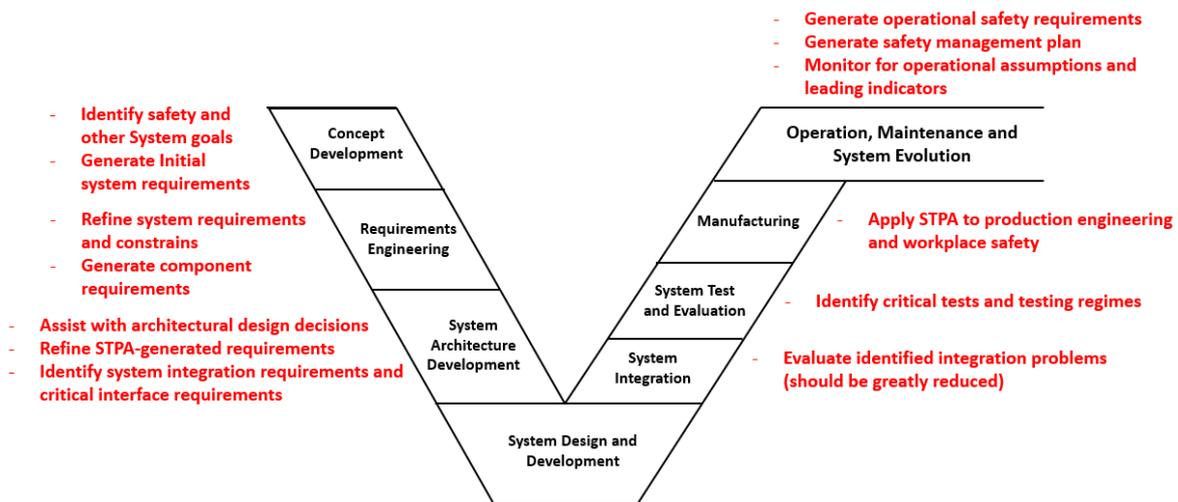
Figure 1: Typical control loop

Figure 2: STPA and System Engineering Process

Although SafetyHAT tool supports STPA methodology, it presents two main differences in the standardized methodology:

1. While STPA methodology consists of two preparatory and two analysis steps, SafetyHAT defines an eight-stage structure.

2. SafetyHAT provides 6 Unsafe Control Action (UCA) types and 26 Causal models Factor Guides (CFG) compared to 4 UCA types and 16 CFGs used traditionally in STPA methodology.

STPA in building risk energy management

The general framework of STPA analysis is applied on a case study within the building energy management sector. A building is considered as an autonomous operating installation and the aim is to reduce the building energy management risks. Based on the scope of the study, potential "accidents" in the building energy ecosystem are analysed. The different levels and corresponding steps for STPA analysis are described.

Level A: System dynamics definition

Step 1: System component import

To identify system-level hazards, the system and its boundaries are initially identified and defined. A system is an abstraction conceived by the analyst; therefore, a decision need to be made about what is included in the system and what the system boundary

is. From the engineering perspective, the most suitable way to define the system boundary for analysis purposes is to include the parts of it within which the system designers have some control.

First, the system foundation is described, by determining individual subsystems. The core subsystem that is examined in this research is the Building Energy Management System.

Building Energy Management Subsystem (BEMS)

The specific subsystem, taken as a reference for the case study, is part of a new holistic methodology under development by the authors for buildings energy efficiency and risk management.

The proposed building energy management model aims to enable the alignment of fine-grain building energy use data to the organizational operational activities (incorporating state of the art business process modelling and annotation techniques). By linking core operational aspects (equipment usage) and environmental conditions (temperature, humidity and luminance) to occupant behaviour underlying business processes and organizational structures, allows for systemic and holistic view over the organizational energy performance. Based on this proposed model, the Building Energy Management System include a number of controllers (e.g., Energy Management System, Presence and Process Module, Comfort Module), actuators (e.g., Smart Home Controller) and sensors (e.g., Temperature/Humidity Sensor, Luminance Sensor, Presence Sensor, Building Meter, Device Status, BPM Tool) to control the basic process of Building Energy Management with the aim of energy demand reduction considering both Business Process hosted in the building and end Users preferences.

External subsystems

As there are cases in which shared control comes from the communication with external subsystems, there is the possibility of conflicting commands by different controllers. Some confusion may result from inconsistencies between the controllers from the external subsystems which may lead to system safety constraints violation and contribute to generation of accident. For the presented case study, the other (external) subsystems (Figure 3) undertake control actions regarding:

- Energy demand reduction in peak periods by load consumption control (Demand Side Management).
- Balance between local building production and consumption to maximize local self-consumption (Local RES)
- Energy storage device operational status optimization (Battery Management)
- Optimized Power charge management

combined with local power consumption (Electric Vehicles).

The system components are presented in Table 1.

Step 2: System connection input

System connections represent information or resource flows or other interactions between subsystems and are indicatively presented in Table2. The interactions between system components are further presented in Table 3. From the analysis, it becomes apparent that energy management modelling at a building level results in a rather complex schema of interconnections.

Step 3: Control action input

A control action is the command issued by a controller that changes the state of the system. Control actions are necessary to ensure the proper system operation and safety. The control actions of the case study are shown in Table 4.

Based on the previous steps and in order to proceed to accident identification, the analytical control structure for BEMS is developed (Figure 4) which is a system model comprising of feedback control loops.
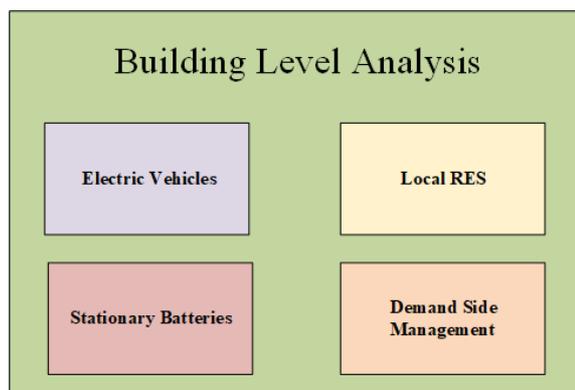

Figure 3: BEMS subsystems

Table 1: System components

| Controllers:	EMS, EV_MS, RES_MS, Battery_MS, Presence & Process Module, Comfort Module, DMS_MS
Actuators	Battery Inverter, RES Inverter, EV Charge Point, Smart Home Devices (HVAC, Lighting, Plug Devices)
Sensors	Temperature/Humidity Sensor, Luminance Sensor, Presence Sensor, Building Meter, Device Status, EV Meter, Battery State, RES Meter/SCADA, Weather Meter, BPM Tool, Aggregate Demand Level
Controlled Process	Building Management
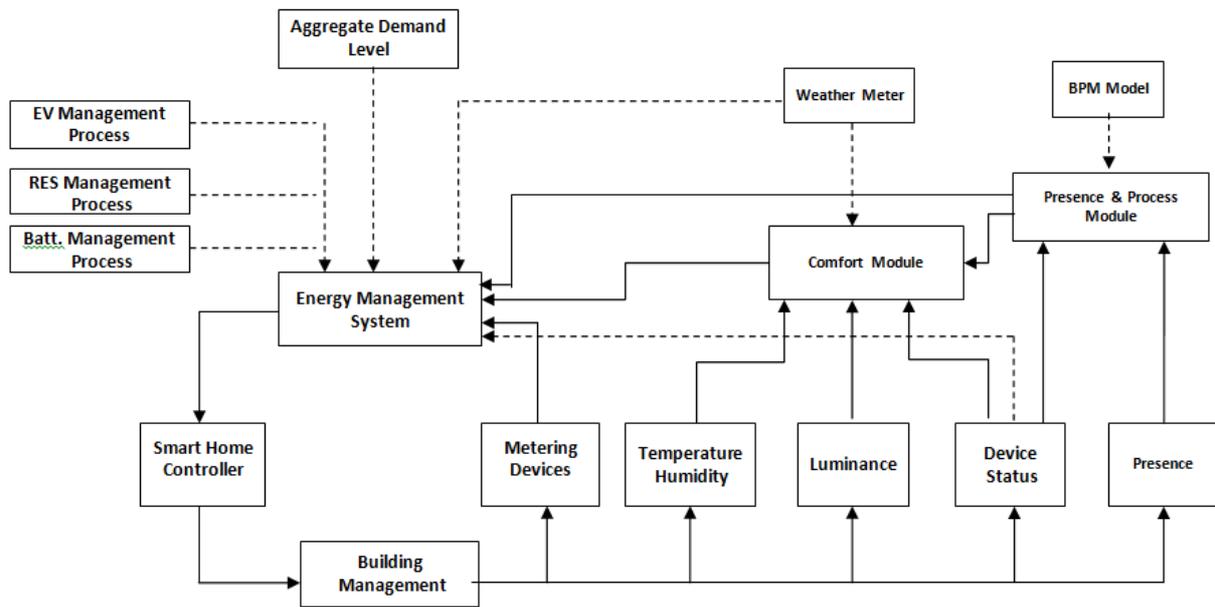
Figure 4: BEMS control structure

Level B: Accidents and Hazard Identification

Step 4: Accident identification

In STPA, an accident is defined as an "unwanted or unplanned event causing loss". Examples of accidents in the context of the building energy management are shown in (Table 5). These are the most common accidents appeared in related bibliography, without being an exhausted list.

Table 2: System connections

Battery Inverter → Battery Management→ Battery State→Battery_MS
EV Charge Point →EV Management → EV Meter →EV_MS
RES Inverter → RES Management → RES Meter/SCADA, Weather Meter→ RES_MS

Table 3: Component connections

Temperature/Humidity, Luminance, Device Status, Presence Sensor/ Weather Meter → EMS → DER Controller → Building Management
Building Meter → EMS → DER Controller→ Building Management
3rd Party Demand Level/ Battery_MS /EV_MS/RES_MS→ EMS → DER Controller → Building Management
Presence Sensor, Device Status→ Presence & Process Module → EMS
Presence, Temperature, Luminance Sensor, Device Status → Comfort Sensing Module → EMS

Table 4: Record of control actions

EV_MS → Deploy EV Charge Mode → EV Charge Point
Battery_MS → Deploy Battery Control → Battery Inverter
RES_MS → Deploy RES Control → RES_Inverter
Presence & Process Module → Deploy Presence Status → EMS, Comfort Module
Comfort Module → Deploy Comfort Status → EMS
EMS → Deploy Building Control (HVAC, Lights) → Smart Home (DER) Controller

Table 5: Building energy management accidents

1. High energy costs due to energy prices
2. Environmental impact: high CO_2 emissions
3. End user dissatisfaction due to comfort and operational conditions
3.1 Excess of non-renewable RES generation)
3.2 High energy cost due to Electric Vehicle Load
3.3 High Battery Management cost

Step 5: Hazards identification

Hazards are defined as system states or conditions that lead to a system accident under a particular set of worst-case environmental conditions. Possible hazards are presented in Table 6. It should be emphasized that the analysis in Level B focuses on risk assessment in the building environment without examining the operational hazards of individual

components (e.g., malfunction of a battery or RES unit as a stand-alone system) which are not being addressed in this section.

Table 6: Building energy management hazards

1. Unacceptable environmental conditions in the building
2. Load misalignment in relation to business processes
3. Peak demand in the local ecosystem (building layout)
4. Local production mismatch with consumption 4.1 Production not available from RES units 4.2 Power not available from a battery pack 4.3 Battery Power Mismatch vs. Consumption 4.4 Non-Controllable Charging Procedure for Electric Vehicles

Level C: Identifying Unsafe Control Actions

After defining the fundamental risk analysis parameters, the next analysis level refers to the 'Unsafe Control Action' (UCA) identification.

Step 6: Identifying UCAs

UCAs assess the evaluation of the potential scenarios that may lead to accidents as those mentioned in the previous step. UCAs are identified by using pre-loaded phrases for control actions, such as:

1. *'Not Provided When Needed'*
2. *'Provided When Not Needed'.*
3. *'Provided Too Early'*
4. *'Provided Too Late'*
5. *'Stopped Too Soon'*
6. *'Applied Too Long'*

For a full analysis, each UCA guidance phrase should be evaluated in each control action. As part of the detailed UCAs analysis, an indicative example is described.

1. A system controller with the corresponding possible control actions is selected, e.g., EMS → Deploy Control HVAC (set point / status / mode).
2. One of the six guidance phrases is selected e.g., NOT PROVIDED WHEN NEEDED
3. One or more system hazards are linked with the UCA description by selecting the risks at the last stage of the process e.g., Deploy Control→NOT PROVIDED WHEN NEEDED →

→ {
   a. Unacceptable environmental conditions (temperature and humidity) in the building
   b. Load misalignment in relation to business processes
   c. Peak Demand in the building
}

Following UCA determining and recording, the next step is to analyse the causal factors.

Step 7: Causal factors analysis

The process of analysing causal factors guides the analysis in identifying how data and connections in the system may lead to unsafe control states. STPA methodology provides guidance for conducting causal factor analysis through the use of guidance factors that provide generic description of how systems or links could lead to unsafe control activities.

Within the methodological framework, the identification of causal factors consists of recording of unsafe control actions by the controller and detecting causes of inadequate feedback from monitors, external systems and / or other controllers. Indicatively, some basic causal factors for Unsafe Control Actions are:

- Decision making inability by controllers
- Incorrect decision-making by controllers
- No control by actuators
- Incorrect control by actuators

The modeling process is described for the test case:

1. Select a controller and the corresponding UCA e.g., EMS → Deploy Control HVAC→ NOT PROVIDED WHEN NEEDED → Cannot be controlled by the Activators.
2. Select a component defined as the causal system e.g., Building Energy Meter.
3. Selection of a causal factor sentence e.g. Sensor INADEQUATE OPERATION with a sample of the details for irregular behaviour.
4. The same process is followed for all connections.

Once the addition of causal parameters for an unsafe description of the control action is done, the modelling process is completed based on the STPA methodology. The relationship 'many to many' captures all possible connections between causes and accidents based on the information we provide in the analysis.

Step 8: Data Export and Safety Recommendations

The result of the analysis is the extensive list of all the individual system components with the recording of unsafe control actions and causal factors. The extracted file from the analysis is a formatted spreadsheet presenting all the individual model co and the final list of individual UCAs (Table 7) and CFs (Table 8) in an overall risk management layout. The number of CFs is evolving on the basis of the UCAs defined in the project. In this context, risk & safety recommendations are developed as suggestions for dealing with hazardous control operations, on the basis of causal factors(those that can lead to an "unsafe state"), and represent possible actions to avoid or reduce the impact of potentially unsafe conditions. Indicative relationships between causal factors and recommendations are depicted in Table 9.

Table 7: Unsafe Control Actions list

COMPONENT_NAME	CONTROL_ACTION	UNSAFE_CONTROL_ACTION	UCA_DESC	HAZARD
Business Module	Deploy Business Constrains	Not provided when needed	Medium Risk - User Preferences violation	Non preference in environmental and operational conditions
Business Module	Deploy Business Constrains	Provided, but the intensity is incorrect (too much or too little)	Low Risk - User Preferences violation	Non preference in environmental and operational conditions
Context Module	Deploy Comfort Constrains	Not provided when needed	Medium Risk - User Preferences violation	Non preference in environmental and operational conditions
Context Module	Deploy Comfort Constrains	Provided when control action is not needed and unsafe	Low Risk - User Preferences violation	Non preference in environmental and operational conditions
DSM_MS	Deploy Building Control (HVAC, Lights)	Provided, but duration is too long or too short	High Risk -Non balance with local RES	Imbalance of local RES with demand
DSM_MS	Deploy Building Control (HVAC, Lights)	Provided, but duration is too long or too short	Low Risk -Peak in Demand	Peak Demand in local network
DSM_MS	Deploy Building Control (HVAC, Lights)	Provided, but executed incorrectly (not executed)	High Risk -Non balance with local RES	Imbalance of local RES with demand
DSM_MS	Deploy Building Control (HVAC, Lights)	Provided, but the starting time is too soon or too late	Low Risk -Peak in Demand	Peak Demand in local network
DSM_MS	Deploy Building Control (HVAC, Lights)	Provided when control action is not needed and unsafe	Low Risk - Imbalance local demand	Non preference in environmental and operational conditions

Table 8: Causal Factors Analysis

CAUSAL_ FACT_NO	CF_DESC	FROM_COMPONENT	TO_COMPONENT
1	Hazardous interaction with other components in the rest of the network	PRESENCE_SENSOR	Business Module
2	Sensor to controller signal inadequate, missing, or delayed: Hardware open, short, missing, intermittent faulty	AGGREGATE DEMAND	DSM_MS
3	Sensor to controller signal inadequate, missing, or delayed: Hardware open, short, missing, intermittent faulty	BATTERY_MS	DSM_MS
4	Sensor to controller signal inadequate, missing, or delayed: Hardware open, short, missing, intermittent faulty	RES_MS	DSM_MS
5	Sensor to controller signal inadequate, missing, or delayed: Hardware open, short, missing, intermittent faulty	WEATHER_METER	DSM_MS
6	Sensor to controller signal inadequate, missing, or delayed: Communication bus error	AGGREGATE DEMAND	DSM_MS
7	Sensor to controller signal inadequate, missing, or delayed: Communication bus error	BATTERY_MS	DSM_MS
8	External disturbances	DSM_MS	BUILDING_ CONTROLLER
9	External disturbances	PRESENCE_SENSOR	Business Module
10	External disturbances	WEATHER_METER	Comfort Module

Table 9: UCAs and associated Recommendations

Controllers	Unsafe Control Actions	Risk Safety Recommendations
DER (HVAC & Light) Command	DSM_Controller does not provide control command when required (Grid constrain, RES/Battery imbalance)	▪ Ensure reliable & uninterrupted DSM software operation ▪ Ensure reliable communication with building actuators and external components (SCADA/ Batteries etc...)
	DSM_Controller provide inefficient control command	▪ Ensure functional reliability of the software: accurate DER model parameters incorporation ▪ Enable training period for definition of model parameters
	DSM_Controller provide inaccurate control command	▪ Ensure functional reliability of interfaces with sensors: energy & environmental sensors reporting accurate contextual conditions ▪ Avoid triggering conflicting actions for load management
Context Module	Context Module does not provide profiles when required	▪ Ensure reliable & uninterrupted Context Profile software operation ▪ Ensure reliable communication of Context Profile module with DSMController
	Context Module provide inefficient command to DSM_Controller	▪ Ensure functional reliability of the software: accurate user preference profiles incorporating contextual parameters ▪ Enable training period for definition of contextual model parameters
	Context Module provide inaccurate results to DSM_Controller	▪ Ensure functional reliability of interfaces with sensors: environmental sensors reporting accurate contextual conditions required for the extraction of contextual profiles ▪ Enable users to adapt their preferences when significant outliers and inaccuracies in the model

Discussion of results

The development and application of the proposed methodology through the corresponding information systems in the screening scenario (energy management in the building sector) has enabled the robustness evaluation of STPA and the reliability of results it provides. The emphasis in this work is to ensure the proper implementation of individual steps of the modelling process and to evaluate the novelty of the proposed approach in comparison to previous methods reported in the literature. The main advancements in this regard are the following:

- Although risk analysis models are being examined in the energy sector, they focus on grid production and operation and not on the consumption segment.

- Risk modelling, in line with demand flexibility, has not been reported in a standardized way in the literature. Demand management is a new mechanism in the energy field with great application opportunity in the framework of smart grids.

- The analysis here is not limited to a single system but it is considered in the context of complex systems and processes. In fact, the emphasis is on the interconnection of individual controllers being independent yet interconnected modules in a building environment.

- STPA provides a direct and clear description of all problem parameters and a thorough analysis of the interactions between them during the design and assessment of the alternative screening scenarios.

Conclusions

This paper presents the adaptation of a new accident analysis technique called STPA (System-Theoretic Process Analysis) and the application of a corresponding computer tool, namely SafetyHAT, to formally develop potential unsafe control situations that may lead to significant hazards in a building energy management system and recommendations for their confrontation.

Considering the complexity of the problem, an attempt to analyse all potential system flaws within a system and check whether they can result in unsafe control actions (UCAs) with traditional risk management models, seems rather unattainable. Instead, it is suggested to start with the hazards, the UCAs and their contextual factors and work backward to identify potential causes. There may still be a large number of causes, but significantly fewer than trying to identify everything that can go wrong and then see if it will lead to a UCA.

Besides the proposed modelling architecture, results from a case study are presented as part of the evaluation of the model feasibility and usefulness for

safety analysis, especially at the early design phase. Results of the evaluation reveal that managing unsafe situations in complex dynamic systems which include a large number of components (controllers, actuators, sensors and controlled processes), could be supported efficiently by the proposed standardized approach.

Our plans for future work consist in deep evaluation of STPA as a hazards identification and analysis methodology with focus on energy applications. Next steps involve a) comparison of the results from STPA with traditional hazard analysis methodologies and further evaluation of results b) further expansion of the methodology to address additional risk and hazards with focus also on smart building environment and smart grids.

Acknowledgments

Nomenclature

CF: Causal Factors
DER: Distributed Energy Resources
DSM: Demand Side Management
EMS: Energy Management System
EV: Electric Vehicles
RES: Renewable Energy Sources
STAMP: Systems-Theoretic Accident Model and Processes
STPA: Systems-Theoretic Process Analysis
UCA: Unsafe Control Action

References

Abdulkhaleq, A. & Wagner, S. (2015). XSTAMPP: An eXtensible STAMP platform as tool support for safety engineering. *The 4th STAMP Workshop 2015*, Boston 10.13140/2.1.3862.0486.

Becker, C. & Van Eikema Hommes, Q. (2014). *Transportation Systems Safety Hazard. Analysis Tool (SafetyHAT) - User Guide (Version 1.0).* Volpe National Transportation Systems Center, U.S. Department of Transportation.

Friedberg, I., Mclaughlin, K., Smith, P., Laverty, D., & Sezer, S. (2017). STPA-SafeSec: Safety and security analysis for cyber-physical systems. *Journal of Information Security and Applications*, 34, pp. 183-196. doi:10.1016/j.jisa.2016.05.008.

Horney, D. (2017). *System-theoretic process analysis and safety-guided design of military systems*. MSc Thesis, MIT, Cambridge, MA.

Krauss, S., Rejzeka M., Hilbesa C. (2015). Tool qualification considerations for tools supporting STPA, *Procedia Engineering*, 128, pp. 15-24.

Leveson, N. (2002). *A new approach to system safety engineering*. Massachusetts Institute of Technology, Cambridge.

Leveson, N. (2004). A new accident model for engineering safer systems. *Journal of Safety Science*, 42, pp. 237–270.

Leveson, N. (2011). *Engineering a safer world: Systems thinking applied to safety*. The MIT Press, Cambridge, MA.

Leveson, N. (2013). *An STPA Primer*. MIT Publication, August 2013.

Ouyang, M., Hong, L., Yu, M. H., & Fei, Q. (2010). STAMP-based analysis on the railway accident and accident spreading: taking the China-Jiaoji railway accident for example. *Safety Science*, 48:5, pp. 544-555.

Wienen, H. C. A., Bukhsh, F. A., Vriezekolk, E., & Wieringa, R.J. (2017). *Accident analysis methods and models - a systematic literature review* (CTIT Technical Report; No.TR-CTIT-17-04). Centre for Telematics and Information Technology (CTIT).