Model-Based Security Engineering for Cyber-Physical Systems: A Systematic Mapping Study

Phu H. Nguyen
Simula Research Laboratory
P.O. Box 134,
1325, Lysaker, Norway
phu@simula.no

Shaukat Ali
Simula Research Laboratory
P.O. Box 134,
1325, Lysaker, Norway
shaukat@simula.no

Tao Yue
Simula Research Laboratory and
University of Oslo
P.O. Box 134,
1325, Lysaker, Norway
tao@simula.no

ABSTRACT

Context: Cyber-physical systems (CPSs) have emerged to be the next generation of engineered systems driving the so-called fourth industrial revolution. CPSs are becoming more complex, open and more prone to security threats, which urges security to be engineered systematically into CPSs. Model-Based Security Engineering (MBSE) could be a key means to tackle this challenge via security by design, abstraction and automation.

Objective: We aim at providing an initial assessment on the state of the art in MBSE for CPSs (MBSE4CPS). Specifically, this work focuses on finding out 1) the publication statistics of MBSE4CPS studies; 2) the characteristics of MBSE4CPS studies; and 3) the open issues of MBSE4CPS research.

Method: We conducted a systematic mapping study (SMS) following a rigorous protocol that was developed based on the state-of-the-art SMS and systematic review guidelines. From thousands of relevant publications, we systematically identified 34 primary MBSE4CPS studies for data extraction and synthesis to answer predefined research questions.

Results: SMS results show that for two recent years (2014-2015) the number of primary MBSE4CPS studies has increased significantly. Within the primary studies, the popularity of using Domain-Specific Languages (DSLs) is comparable with the use of the standardized UML modeling notation. Most primary studies do not explicitly address specific security concerns (e.g., confidentiality, integrity) but rather focus on security analyses in general on threats, attacks or vulnerabilities. Few primary studies propose to engineer security solutions for CPSs. Many focus on the early stages of development lifecycle such as security requirement engineering or analysis.

Conclusion: The SMS does not only provide the state of the art in MBSE4CPS, but also points out several open issues that would deserve more investigation, e.g., the lack of engineering security solutions for CPSs, limited tool support, too few industrial case studies, and the challenge of bridging DSLs in engineering secure CPSs.

Keywords

Cyber-Physical Systems; Security; Model-Based Engineering; Security Engineering; Systematic Mapping; Snowballing; Survey

1. INTRODUCTION

Nowadays, Cyber-Physical Systems (CPSs) could be considered as the game changer in a wide range of industries (e.g., manufacturing, energy, healthcare and automotive industry), infrastructures (e.g., transportation, water management, oil and gas pipelines, wind farms), facilities (e.g., airports, space stations and buildings), and military (e.g., drones and unmanned aerial vehicles). As stated in [53], "*cyber-physical systems (CPSs) are physical and engineered systems whose operations are monitored, coordinated, controlled and integrated by a computing and communication core*". An example of CPSs is seen in modern power grid systems. In such a smart grid system, information and communication technology (ICT) is increasingly integrated throughout the grid to support novel communication and control functions among physical resources such as wind farm, solar farm, smart meters and information and control systems. Data (e.g., meter readings) collected from the sensors of physical resources (e.g., smart meters) are transmitted to information and control systems for live monitor and control (e.g., remote disconnect of smart meters). Computations based on these two-way communications allow the most efficient utilization of renewable resources, and the great customization of smart grid services. CPS technology would be expected to transform the way people interact with engineered systems like the Internet has transformed the way people interact with information [1].

The more human beings surrounded by CPSs, the more important that these CPSs must be secure. A single security issue in smart grid could lead to city blackout or even country blackout. Large scale attacks in the software side of highly specialized industrial control systems were supposed to be very unlikely. However, the Stuxnet worm attack in the summer of 2010 was a wakeup call on the security of industrial CPSs [31]. By interfering the software that controls physical devices in a nuclear power plant, Stuxnet worm could destroy those physical devices or even the power plant. Stuxnet proved that even isolated industrial CPSs could be compromised, causing them to have unexpected (physical) operations, e.g., self-destruction. Moreover, many modern CPSs would unavoidably need to connect to the Internet that could bring much more security challenges. The security of CPSs is of paramount importance also because in many cases security could mean the physical safety of human beings around these systems. Put aside industrial systems, one of the biggest cyber-security threats in 2016 was predicted to come from hacked medical devices [23]. By hijacking insulin pumps and pacemakers that are part of CPSs in healthcare domain, hackers could hold patient's life ransom as warned in [23]. Again, this kind of threat urges the security of CPSs to be taken into account very early, seriously, and systematically. An important lesson should be learned from the way information systems had been engineered in the past is that security often came as an afterthought [17]. If security is not taken into account very early in the development lifecycle, it is nearly impossible to engineer security requirements properly into any complex system. One of the main reasons is that security

requirements are often scattered and tangled throughout system functional requirements. Therefore, the security of CPSs should be engineered "by design" early in the CPSs' development.

However, CPSs are in many cases highly complex and making sure of their security is very challenging. Besides the cyber security challenges of CPSs, the security of the physical parts of CPSs, which are controlled by software-defined controllers based on computational algorithms, is indeed a new critical challenge. For example, physical devices like smart meters are deployed at the "client side", where hackers could have better chance to tamper them and intrude into smart grid. Software is the soul of CPSs. Therefore, innovative, sound software security engineering methodologies are sought to address the security challenges of CPSs. Some researchers consider Model-Based Engineering (MBE) or Model-Driven Engineering (MDE) as one of the key solutions to the handling of complex systems [9], including CPSs [6]. One of the main ideas of MBE/MDE is the engineering at the model level, a higher level of abstraction than the code level. This would allow better engineering security together with the system as well as providing the foundations for (semi-) automated (formal) verification or validation of the security of complex systems. Indeed, MDE methods have been actively developed for engineering the security of complex software systems very early and throughout the development life cycle as surveyed in [47]. In a recent study that assessed the state of the art and the state of the practice in the verification and validation of CPSs, the authors suggest that "*model-based approaches are gaining momentum, and it seems inevitable that model-based approaches will emerge that can be applied to general purpose CPSs*" [76]. By engineering systems via computer-readable models, model-based security engineering (MBSE) techniques could provide solutions to address the challenges for the security of CPSs. We call the MBSE approaches that are specifically developed or adopted for CPSs as MBSE4CPS. However, it remains a big question on how extensively the MBSE4CPS approaches have been developed. This paper aims to give an answer to this question.

After conducting a trial survey on the topic of MBSE4CPS, we found that this is an emerging interdisciplinary research area among several research fields such as software (system) engineering, (software) security engineering, and electrical/system engineering. Therefore, a systematic mapping study (SMS) would be useful to provide a picture of the MBSE4CPS research so far, in the interests of researchers and practitioners in the research fields mentioned above. We followed the latest guidelines in [52] to conduct a SMS on the existing primary MBSE4CPS studies. Thousands of relevant papers have been systematically filtered from four main online publication databases, and from an extensive snowballing process [71] to finally obtain a set of 34 primary MBSE4CPS studies. We extracted and synthesized data from the primary MBSE4CPS studies to answer our research questions. In the end, the key contributions of this work are our answers to the following research questions (as well as their sub-questions in Section 3):

RQ1 What are the publication statistics of the existing primary MBSE4CPS studies in the literature?
RQ2 What are the existing primary MBSE4CPS studies and their characteristics?
RQ3 What are the open issues of MBSE4CPS research?

Besides, it is important to note that in complex systems such as CPSs, uncertainty is very likely to happen and must be handled [3]. From security's point of view, uncertainty in CPSs could lead to serious security issues. For example, some uncertainties in the functionalities of CPSs might lead to vulnerabilities that could be exploited by an adversary, either attacker or malicious user. Vice versa, any uncertainty in the specification, implementation and evolution of security mechanisms might cause other uncertainties in the functionalities of CPSs, e.g., incorrect access control can disable some physical processes, especially whose real time requirement is critical. On the other hand, security attacks could also cause uncertainties in the functionalities of CPSs. Therefore, while conducing this SMS we did keep in mind to check if any primary MBSE4CPS study explicitly deals with uncertainty.

The remainder of this paper is structured as follows. Section 2 provides some background concepts that are used in this paper. Then, we present in Section 3 our approach to conduct this SMS. Section 4 contains our classification schemes for the primary MBSE4CPS studies and other criteria for supporting the data extraction and comparison among these primary studies. Key results are described in Section 5 followed by threats to validity in Section 6. Related work is presented in Section 7. Finally, Section 8 concludes the paper with the major findings and some directions for future work.

2. BACKGROUND

In this section, we provide some background concepts that are used throughout this paper. First, we recall in Section 2.1 the definition of SMS in relation to other types of secondary study such as Systematic Literature Review. In Section 2.2, the scope in which an approach can be considered as an MBSE approach is discussed in comparison with related concepts such as Model-Driven Security (MDS). Then, in Section 2.3 we define the scope in which a system can be considered as a CPS, and some fundamental security concepts in the context of CPSs.

2.1 Systematic Mapping Study vs. Systematic Literature Review

According to [34], there are three different kinds of *secondary study* that would complement each other: *Systematic Literature Review (SLR*), *SMS*, and *Tertiary Review (TR)*.

- Secondary study: *"a study that reviews all the primary studies relating to a specific research question with the aim of integrating/synthesizing evidence related to a specific research question [34]."*
- SLR: *"A form of secondary study that uses a well-defined methodology to identify, analyze and interpret all available evidence related to a specific research question in a way that is unbiased and (to a degree) repeatable [34]."*
- SMS: *"A broad review of primary studies in a specific topic area that aims to identify what evidence is available on the topic [34]."*
- TR: "*A review of secondary studies related to the same research question [34]*."

As can be seen from [52] and [33], SMS and SLR may have similarities in conducting some first steps such as primary studies search and selection. However, their goals as well as their approaches to data analysis are different [52]. SMS aims to discover research trends with general research questions for classification and aggregation of relevant studies according to predefined (high level) categories, e.g., publication trends of a research domain over time. In a SMS, the evidence in a domain is plotted at a high level of granularity [34]. SLR on the other hand focuses on more (low lever) detailed aggregated evidence in terms of the research outcomes driven by very specific research questions, e.g., whether a methodology is practically useful by

industry [52]. More details on the differences between SMSs and SLRs can be found in [52], [34].

In a SLR as well as a SMS, the search and selection process of primary studies must be transparent and exhaustive to identify as many relevant research papers as possible in the focus of the review. Database search on online publication repositories such as IEEE Xplore[1] is so far the most popular search strategy employed by secondary studies [52]. However, database search still has some limitations such as the construction of search strings and limited support by search engines. Therefore, the snowballing search strategy has been introduced in [71] that could complement database search as both of these search strategies were employed in the SLR [47]. The snowballing search strategy consists of the following main steps: 1) identify a starting set of primary papers (e.g., by using database search); 2) identify further primary papers using the list of references in each primary paper (backward snowballing); 3) identify further primary papers that cite the primary papers, e.g., by using Google Scholar[2] (forward snowballing); 4) (recursively) repeat Steps 2 and 3 until no new primary papers are found. In this SMS, we employed both database search and snowballing.

2.2 Model-Based Security Engineering

2.2.1 MBE and MDE

MBE could be the key to engineer complex systems, including CPSs and their security. By modeling the desired system and manipulating models, the level of abstraction is higher than code-level that brings several significant benefits, especially regarding security engineering. First, security concerns (e.g. confidentiality, integrity, availability) can be considered together with the business logic (and other quality attributes like performance) very early, which is crucial in engineering secure systems. As found out in [47], domain-specific languages (DSLs) are normally developed and used in security engineering because of their expressiveness ability for capturing security mechanisms. In other words, a DSL that is tailored for specifying a specific security aspect (e.g., access control) should be more expressive than a general modeling language like UML. However, the UML profile mechanism can be used for the definition of security-oriented DSLs as surveyed in [47]. Besides UML profiles, some other approaches surveyed in [47] introduced non-UML based DSLs.

Second, reasoning about the desired systems at the model level could enable model-based verification and validation methods with tool support, which are important for security analysis. If transforming security models into possible inputs for formal methods (and existing tools, e.g., Alloy [41]) is feasible, formal methods such as model checking could be employed for verifying security properties. Model-based security testing methods could be employed for validating the resulting secure systems (especially in where formal methods would not be applicable).

Third, engineering at model level would enable automation provided by automated model-to-model transformations (MMTs) and model-to-text transformations (MTTs). MMTs can take part in the key steps of engineering process, e.g. for composing security models into business models, or transforming models between different DSLs. MTTs can be used for generating code, including security mechanisms, e.g., a configured access control mechanism. The automation would make the development process more productive with higher quality compared to a hand-written code development process [69].

To set the scope of what can be considered as an MBSE approach (and then MBSE4CPS), we recall the concepts MBE, MDE, and Model-Driven Development (MDD) from [12]. According to [12], models in MBE approaches may not necessarily be the central artifacts in the development lifecycle. For example, models in an MBE approach may be used for either documentation or verification purposes, but may not necessarily or possibly be used for implementation. On the other hand, models in MDE approaches are primary artifacts that "drive" the development, evolution, or migration tasks [12]. If an MDE approach only focuses on development, it is called MDD. Therefore, MDD is a subset of MDE. Similarly, MDE is a subset of MBE as discussed in [12] because models in MDE must be primary "*driving*" artifacts and cannot just be for documentation purpose or any single engineering purpose as in MBE scope.

2.2.2 MBSE and MDS

In [47], a concrete definition and scope of MDS has been given. Roughly speaking, MDS is a subset of MDE in which secure systems are the focus of engineering. Similarly, MBSE is a subset of MBE. Because CPSs are the new generation of engineered systems, security-engineering approaches based on models have just emerged. In this paper, we are interested in the broad sense of security-engineering approaches based on models, i.e. MBSE. In developing secure systems, MBSE could play an important role, e.g., in the verification and validation of secure systems regarding their security properties. Models in an MBSE approach may be used for design or implementation purposes but also may only be for security analysis, or verification and validation purposes. MBSE approaches that are developed specifically for CPSs are called MBSE4CPS.

2.3 Cyber-Physical Systems and Security

2.3.1 Cyber-Physical Systems

"*Cyber-physical systems (CPS) are engineered systems that are built from, and depend upon, the seamless integration of computational algorithms and physical components*" [1]. According to [53], "*CPSs are physical and engineered systems whose operations are monitored, coordinated, controlled and integrated by a computing and communication core*". We used these definitions to search for publications in CPSs' application domains and relevant domains such as embedded systems, system of systems. More specifically, we did take into account also embedded systems or the systems of systems that have CPSs' characteristics.

Based on the definitions of CPSs above, many modern systems in different domains can be classified as CPSs. In [32], the popular application domains of CPSs have been surveyed and are summarized as follows: Vehicular Systems and Transportation (e.g. smart car); Medical and Health Care Systems; Smart Homes and Buildings; Social Network and Gaming; Power and Thermal Management; Data Centers (operating like CPSs to keep energy costs for computation and cooling minimal); Electric Power Grid and Energy Systems (e.g. smart grid); Networking Systems; Surveillance.

2.3.2 The Security of CPSs

The security concerns (objectives) of CPSs are not different from the traditional security concerns of computer security, e.g., confidentiality, integrity, availability (CIA), and accountability. Only that the details of each security concern must be interpreted in the context of CPSs, e.g., as given in [13] or [49], which bring

[1] IEEE Xplore, http://ieeexplore.ieee.org/
[2] Google Scholar, https://scholar.google.com

up new security challenges, e.g., in protecting physical devices. In this paper, we refer to security terms described in [39] such as security threats, vulnerabilities, attacks, and security solutions as different aspects (*security aspects*) to be considered while engineering security. On the other hand, *security concerns* refer to security objectives (e.g., CIA, Accountability) and mechanisms (e.g., Authentication, Authorization, Encryption). *Security solutions* are the combination of security mechanisms according to security objectives to mitigate security vulnerabilities. We adopt some definitions of the generic security concerns from [11, 39] and CPS specific ones from [13] as follows.

"*Confidentiality is the concealment of information or resources*" [11]. Unauthorized parties are prevented from knowing the information or resources, even from being aware of their existence. In CPSs, the state of the physical system must be kept confidential from unauthorized parties, i.e. sufficient security mechanisms must prevent eavesdropping on the communication channels, e.g. between a sensor and a controller, and between a controller and an actuator. Moreover, in some CPSs that have sensitive users' data, these data must be protected from unauthorized access.

"*Integrity refers to the trustworthiness of data or resources, and it is usually phrased in terms of preventing improper or unauthorized change*" [11]. Integrity in CPSs can be viewed as the ability to maintain the operational goals by preventing, detecting, or surviving deception attacks in the information sent and received by sensors, controllers, and actuators. If integrity is not ensured, deception could happen, i.e., "*when an authorized party receives false data and believes it to be true*" [13].

"*Availability refers to the ability to use the information or resource desired*" [11]. Lack of availability could result in denial of service (DoS). A DoS attack is characterized by an explicit attempt to "*prevent the legitimate use of a service*" [38]. The goal of availability in CPSs is therefore, to maintain the operational goals by preventing or surviving DoS attacks to the information collected by the sensor networks, commands given by controllers, and physical actions taken by actuators. There could be new challenges for ensuring availability in many CPSs whose real-time constraints are critical.

Accountability: Besides CIA, accountability is another security concern that is also important in many applications. Accountability refers to the ability to keep track of who did what and when.

In any CPS, efficient control over some physical processes is the main goal. Therefore, information integrity and availability are vital to ensure that a control state closely mirrors a physical system state. Cryptography, access control, and authentication are some security mechanisms that could provide integrity in systems. However, any security mechanism employed must also provide sufficient availability. This constraint often limits the utilization of security mechanisms because they may deny access to a critical function [57]. The insufficient interaction between security mechanisms and CPSs' operations could cause uncertainty in CPSs. For example, an inadequate access control mechanism could block or slow down access to a physical device whose real-time requirements are critical.

2.3.3 CPS Uncertainty and Security

We recall a definition of uncertainty from [3]: "*Uncertainty is a state of a CPS that is unpredictable, a future outcome from the state may not be determined, or there is a possibility of more than one outcome from the state*". Uncertainty and security are two of the main essential characteristics of CPSs bringing huge challenges that need to be addressed in research [28]. Uncertainty and security of CPSs could intertwine in different ways. A security incident (e.g., caused by attackers) or misconfiguration may lead to uncertainty. Vice versa, uncertainty may lead to security vulnerabilities that could be exploited by attackers. This security-related uncertainty can occur in a CPS because of 1) ambiguous or missing security requirements; false security assumption; false security goals; 2) the possible security misconfiguration, incorrect implementation, or wrong security policy that could prevent the CPS to operate certainly; and 3) the possible security vulnerabilities or misconfiguration of the CPS that could lead to successful security attacks; the unpredictable security attacks aiming at the CPS.

3. SYSTEMATIC MAPPING APPROACH

We conducted our SMS by following the latest systematic mapping study guidelines [52] as well as consulting other relevant guidelines and studies reported in [71], [33], [10] and [34] for example. Based on our research questions (Section 3.1), we identified search terms (Section 3.2) and designed a search strategy (Section 3.4) to find the primary studies that can answer our research questions. It is also important to clarify the inclusion and exclusion criteria (Section 3.3) to reduce possible bias in the selection process (Section 3.4). The process of data extraction and synthesis of the primary studies was based on a set of evaluation criteria (Section 4).

3.1 Research Questions

To answer our general research questions raised in Section 1, we detail them into sub-questions. As discussed in [34], the research questions of a SMS are normally generic and related to research trends, e.g., to find out which researchers, how much activity, etc. To be more specific on what publication statistics we want to find out, the RQ1 is divided into four sub-questions.

First, we are interested in the trend of the primary MBSE4CPS studies published over time per year. ***RQ1.1*** *- In which years were the* primary *MBSE4CPS studies published and what is the annual number of publications?* Answering RQ1.1 would allow us to discover when the first primary MBSE4CPS study was published and the frequency of the publication of the primary MBSE4CPS studies. We could base on this finding to assess if this research topic has been getting more attention from research community.

Second, we would like to know relatively about the publication venues of the primary MBSE4CPS studies, e.g. if a publication venue is a journal, conference, or workshop. The primary MBSE4CPS studies are the approaches that develop or leverage model-based software security engineering techniques for CPSs. Therefore, these studies could be published at different kinds of venues such as software engineering venues, security-engineering venues, or system engineering venues. ***RQ1.2*** *- In which targeted venues (e.g., software engineering venue, security engineering venue), and venue types (e.g., conference, journal, workshop) were the primary MBSE4CPS studies published?* Note that there has not been yet any specialized MBSE4CPS journals or conferences. Answering RQ1.2 would enable us to know which venues have been the targets for publication of primary MBSE4CPS studies. The venue types could also provide some hints on the maturity of the primary MBSE4CPS studies, i.e., papers published at journals are supposed to report more mature studies than papers published at conferences and workshops.

Third, the involvement of industry in MBSE4CPS studies would be an indicator of industry's interest of MBSE4CPS topic as well as the research collaboration among industry and academia. Therefore, we want to know whether the authors of the primary MBSE4CPS studies work in academia or industry. A paper is classified as *academia* if all authors come from academy (university or research institute), *industry* if all authors come from a company, and *both* (academy and industry) if there is a mix of authors from academy and company. ***RQ1.3*** *- What is the distribution of publications in terms of academic and industrial affiliation?*

Fourth, we would like to know in which country that the primary MBSE4CPS studies have been researched. ***RQ1.4*** *- What is the geographic distribution of the research on MBSE4CPS?* Answering RQ1.4 would allow us to identify which countries (or continents) are leading in terms of research publications in this domain. The findings could be related to the research focuses on CPSs that have been being promoted by many countries such as the United States and in the European Union (EU) [25].

To be more specific on what characteristics of the primary MBSE4CPS studies we want to examine, the RQ2 is divided into seven sub-questions.

It is important to understand what security concerns are addressed in each primary MBSE4CPS study. From security engineering point of view, security approaches must be driven by concrete security concerns. ***RQ2.1*** *- What security* concerns *(e.g., confidentiality, integrity, availability) were addressed in the primary MBSE4CPS studies?*

Each security engineering approach could focus on solely or in combination on different security aspects such as attacks, or threats, or vulnerabilities, or solutions. For each *primary* MBSE4CPS study, we want to know exactly which aspects are mainly tackled. ***RQ2.2*** *- Which security aspects (e.g., attack, threat, vulnerability, solution) were focused on?*

In any primary MBSE4CPS study, security aspects should be modeled or specified. These models are then engineered and/or transformed at the development processes of CPSs. ***RQ2.3*** *- How were the security aspects modeled (specified) and engineered (transformed)?*

As any software engineering approach, each primary MBSE4CPS study could focus on supporting specific engineering phases in the development lifecycle. ***RQ2.4*** *- Which engineering phases that the primary MBSE4CPS studies focused on or supported, e.g., requirement engineering, design, and testing? Do the approaches report tools?*

Similar to any software engineering approach, we could use the research contribution types and research types as discussed in [52] to analyze the primary MBSE4CPS studies. ***RQ2.5*** *- What types of contributions (e.g., process, tool, method) and what fine-grained types of research (e.g., opinion, conceptual, solution, validation, evaluation) were the primary MBSE4CPS studies?*

From CPSs perspective, we want to know what kinds of CPSs that the primary MBSE4CPS studies applied for and whether to real cases? ***RQ2.6*** *- What CPSs were these primary MBSE4CPS studies applied for? What kinds of case studies (academic or industrial) were used to evaluate the approaches?*

As mentioned in introduction, uncertainty would need to be specifically tackled for CPSs. We want to examine if any primary MBSE4CPS study has proposed to deal with uncertainty. ***RQ2.7*** *- Has any primary MBSE4CPS study dealt with uncertainty?*

RQ3 is divided into two sub-questions. Based on the characteristics of the primary MBSE4CPS studies, we want to find out the open issues that would deserve more investigation in the future and some potential directions to tackle these issues. ***RQ3.1*** *- What are the open issues of MBSE4CPS research?* ***RQ3.2*** *- What research directions could be recommended for tackling the open issues?*

3.2 Search String

From the research questions, we identified the search terms and grouped them into four groups: population, intervention, comparison, and outcome (PICO) [33].

The population terms are the keywords that represent the CPSs domain. We used the keywords of the most dominant application domains of CPS technology, e.g. smart grid.

- Population: ("cyber-physical system" OR CPS OR "smart grid" OR "power grid" OR "smart car" OR "automotive cyber-physical system" OR "pervasive healthcare system" OR "unmanned aircraft system")

The intervention terms are the keywords that represent the MBE techniques.

- Intervention: (model OR modeling OR model-based OR model-driven)

The comparison terms represent the security concerns or aspects. These are the key terms in security engineering as presented in [39]. Besides security terms, we also included a specific keyword "uncertainty".

- Comparison: (security OR confidentiality OR integrity OR availability OR accountability OR authentication OR authorization OR "access control" OR attack OR threat OR vulnerability OR uncertainty)

The outcome terms represent the goals of engineering process.

- Outcome: (architecture OR design OR verification OR validation OR test OR analysis)

To form the search string, we used the conjunction of the groups of terms above, i.e., Population (group) AND Intervention AND Comparison AND Outcome. The search string was the input for our database search process described in Section 3.4.

3.3 Inclusion and Exclusion Criteria

The aim of this SMS was to identify and classify papers related to MBSE approaches for CPSs. The inclusion criteria (IC) were:

- (IC1) The paper must have an MBSE context. This means that model(s) have to be used in some security engineering processes.
- (IC2) The paper must address cyber security.
- (IC3) The paper must aim at CPSs, either in general or in a specific application domain of CPSs such as smart grid.

We excluded papers that met at least one of the following exclusion criteria (EC):

- (EC1) Papers not addressing cyber security are excluded.
- (EC2) Papers not proposing MBE approach are excluded.
- (EC3) Papers not addressing CPSs are excluded.
- (EC4) Grey literature and non-English papers are excluded.

- (EC5) Non peer-reviewed papers, keynotes, workshop reports, books, theses, and dissertations are excluded.
- (EC6) Any incomplete or old version of a publication was excluded. For example, we excluded some workshop or conference papers once we had found the extended journal versions of those papers.

3.4 Search Strategy and Selection Process

According to [52], database search via online databases such as IEEE Xplore is the most common way of finding primary studies for a SMS or SLR. Besides, by searching on different databases, we could have more chances to find papers related to MBSE4CPS from different research communities. We expected that researchers working on CPSs and security could be from different research areas such as electrical engineering, software engineering, security engineering. Moreover, to overcome some limitations of database search as pointed out by [72], we employed the snowballing strategy [71] for complementing the set of primary studies found from the database search. Therefore, our search and selection process consists of two phases as follows.

3.4.1 Database search

We used (with adaptation if necessary) the search string above on four online databases: IEEE Xplore, ACM DL[3], Scopus[4], and Springer Link[5]. The main reasons of using these databases are because these are big and common databases; and they (except Springer) allow the search results to be exported in a format that can be directly imported into EndNote tool[6]. We used EndNote to manage the candidate papers in our selection process. EndNote tool also allowed removing duplicates in the candidate papers easily. We searched for papers in the range from 2001 until 2015 because the earliest MBSE studies were only found in the early twenty-first century [37].

Step 1. *Preprocessing*: Based on the search results returned from search engines, we merged them to eliminate duplicates with tool support of EndNote (Step 1 in Figure 1). We also manually removed the books, white papers, tables of contents, etc. Figure 2 shows the distribution of aggregated search results from four databases, per year. As can be seen in Figure 2, the number of related papers found by search engines sharply increased from 2001 to 2015.

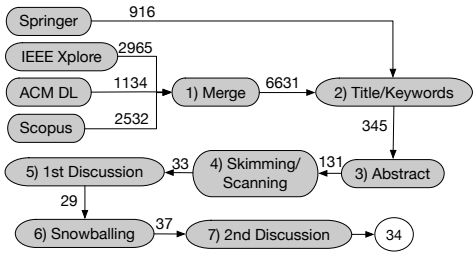

Figure 1. Primary studies selection process

Steps 2, 3, 4. *Multilevel of contents checking*: From the set of candidate papers, we filtered out the primary MBSE4CPS studies according to the predefined inclusion/exclusion criteria. Our selection process was based on multi levels of checking: title, abstract, and skimming, scanning through the main contents of each candidate paper. To be more specific, for each candidate paper we first read the paper's title, keywords to see if it could be decided on the IC and EC. If the title and keywords are insufficient for us to decide to include it for the next round or exclude it immediately, we further checked the paper's abstract. If we still cannot have an inclusion or exclusion decision based on the abstract, we further skimmed/scanned the paper's full content. Besides EndNote, we used Mendeley[7] tool to manage the papers whose detailed contents needed to be reviewed (by skimming and scanning). Note that we rather kept any candidate paper in doubt at any point for further checks later. In the end, we still had to arrange discussion among reviewers to crosscheck the candidate papers in doubt and agreed on final decisions to include or exclude them.

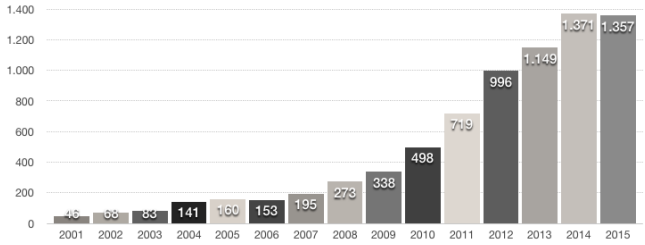

Figure 2. The distribution per year of aggregated search results from four databases

Step 5. *Crosschecking and face-to-face discussion #1:* Borderline papers were discussed among the authors of this paper to reach inclusion/exclusion decisions. In the end, we obtained a set of 29 primary papers from database search as showed in Figure 1.

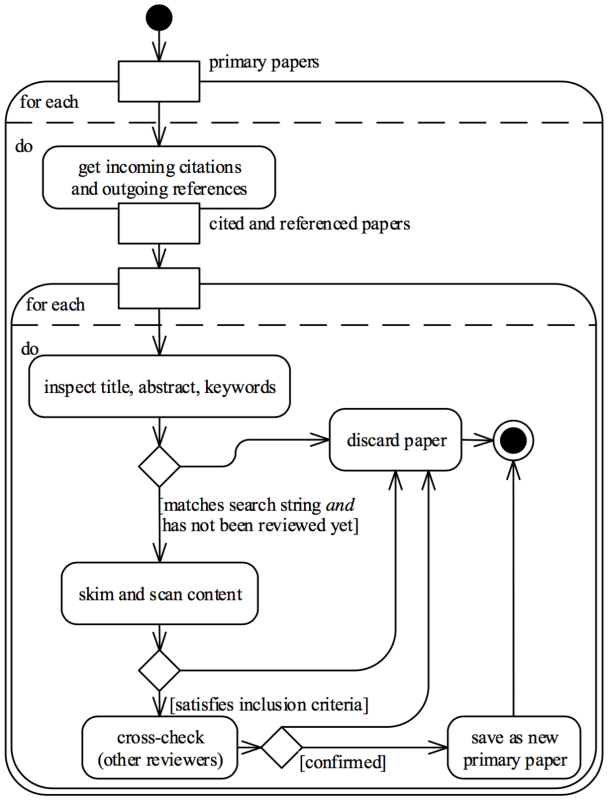

Figure 3. Our selection process while snowballing (adopted from [47])

[3] ACM Digital Library, http://dl.acm.org
[4] Scopus, http://www.scopus.com
[5] Springer Link, http://link.springer.com
[6] Endnote, http://endnote.com
[7] Mendeley, http://mendeley.com

3.4.2 Snowballing search

As pointed out by [72] and based on our own experience from [47], we conducted a secondary search process to overcome some limitations of database search by using the snowballing strategy [71] on the selected primary papers obtained after the database search.

Step 6. *Snowballing:* This means that we examined the list of references and citations (from Google Scholar) of each primary paper obtained in the first step to find new primary papers (see Figure 3). For each paper in the set of cited and referenced papers of 29 primary papers above, our selection process was again based on multi levels of checking: title, abstract, and skimming, scanning through the main contents. The snowballing process was also applied recursively to the newly found primary papers. We found out eight more candidate papers from this snowballing process.

Step 7. *Crosschecking and face-to-face discussion #2:* After discussion on some borderline papers, we excluded three of them. In total, we obtained a set of 34 primary MBSE4CPS studies as showed in Figure 1.

4. CLASSIFICATION SCHEMES

To analyze the primary MBSE4CPS studies for answering our research questions, we defined four categories of classification criteria. As it can be seen in Figure 4, our classification schemes are based on the main artifacts of MBE, security engineering, and CPSs, plus some general classification artifacts for research publications. More specifically, we included in our classification schemes the key artifacts that are selected from the evaluation taxonomy of MDS in [47], from the key security concepts in [39], from the Microsoft Security Development Lifecycle (SDL) [40], and from the application domains of CPSs in [32]. In addition, we also use some general classification artifacts in terms of research contribution type and research type as discussed in [52] to classify the primary MBSE4CPS studies.

From MBE perspective, we would like to know which ***modeling notation(s)*** have been used in the primary MBSE4CPS studies. Modeling notation is important to specify and capture the domain knowledge for engineering purposes. UML-based modeling notation is standard, but domain-specific (modeling) languages have also been introduced for engineering secure systems [47]. Apart from the modeling notation, ***modeling methodology*** also plays a big role in MBSE. Aspect-oriented modeling (AOM) methodology [24, 68] is supposed to provide advantages in specifying crosscutting properties of systems like security. We would like to check if AOM has been leveraged in MBSE4CPS. Besides, how security aspects and system elements of CPSs are specified depends on what ***kinds of model*** have been used in the primary MBSE4CPS studies, i.e., UML-based structural models (e.g., class diagrams, composite structure diagrams), UML-based behavioral models (e.g., sequence diagrams, state diagrams), or domain-specific models (DSMs created by DSLs, e.g., Security Analysis Language [19]). Another important artifact of MBE is ***model transformations*** which could be considered as the heart and soul of model-driven software development [56]. During model-based engineering processes, model-to-model transformations (MMTs) can be used for different engineering purposes such as composing security models with system models, or transforming secure design models to some types of models that can be used for security analysis. MMTs can also be classified as endogenous MMTs (between models expressed in the same language) or exogenous MMTs (between models expressed using different languages). On the other hand, model-to-text transformations (MTTs or code generation techniques) can be used for generating implementation code, including security configurations.

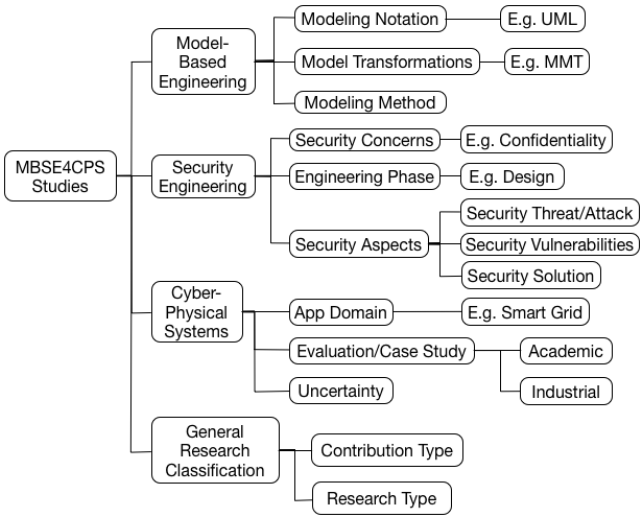

Figure 4. Our classification schemes of MBSE4CPS studies

From security engineering perspective, we would like to examine which ***security concerns*** have been focused on, e.g., confidentiality, integrity, availability, accountability, authentication, and authorization (CIAAAA). Besides, security-engineering approaches would focus on some specific **security aspects**, e.g., attacks, threats, vulnerabilities, or solutions. We recall some definitions from [39] as follows: "***Threat*** *is a potential violation of security.* ***Attack*** *is an action that could cause a violation of security to occur.* ***Vulnerability*** *is a weakness of an asset or control, which may be exploited by a threat.*" ***Security solutions*** are the mitigation of security vulnerabilities. For each primary MBSE4CPS study, we were also interested in knowing which security ***engineering phase***(s) that the approach focused on. To have a unified view on the security engineering phases, we based our classification on the main stages of the SDL [40], i.e., requirements, design, implementation, verification, release, and response. We can see that this SDL is relatively similar to the software development life cycles. As stated in the RQ 2.4, we want to examine if any approaches have ***tool support*** for the security engineering phases. For each approach, tool support can be classified as: new tool developed, existing tool adopted, or no tool support. For each tool, we check tool platform, tool input and tool output.

From CPSs perspective, we wanted to know what kinds of CPSs and their security were the focuses of the primary MBSE4CPS studies. This information would point out the CPSs' application domains, which have attracted the attention from MBSE4CPS research community. We adopted the ***application domains*** of CPSs surveyed in [32]. To evaluate the involvement of industry (via real-world case studies) in this research area, we would like to know if the CPSs used as ***case studies*** in the primary MBSE4CPS studies are from industry or academia. Moreover, for our interest in the ***uncertainty*** of CPSs as mentioned earlier, while conducing the SMS we also kept in mind to check if any primary MBSE4CPS study explicitly deals with uncertainty.

From general research classification, besides the specific artifacts of MBSE4CPS, we also used the general classification artifacts in terms of research ***contribution type*** and ***research type*** as discussed in [52] to classify studies. The research contribution

types are: method (techniques/approaches), model, metrics, tools, and open items (identified issues to be addressed). The classification of research types is recalled from [70] in Table 1.

Table 1. Research type classification [70]

Category	Description
Validation research	*"Investigating a proposed solution, which is novel and has not yet been implemented in practice. Investigations are carried out systematically, i.e., prototyping, simulation, experiments, mathematical systematic analysis and mathematical proof of properties."*
Evaluation research	*"Evaluating a problem or an implemented solution in practice, i.e., case studies, field studies and field experiments."*
Proposal of solution	*"A novel solution for a problem or new significant extension to an existing technique."*
Conceptual proposal	*"A new way of looking at things by structuring in form of a conceptual framework or taxonomy"*
Opinion paper	*"The author's opinion on whether a certain technique is good or bad"*
Experience paper	*"Personal experience of the author, i.e., what and how something has been done in practice."*

5. RESULTS

The first author of the paper used Microsoft Excel spreadsheets to record data extracted from the primary MBSE4CPS studies. Several revisions of the spreadsheets were made afterwards while extracting data to better support the extraction process and enable the comparability between studies. After synthesizing the data, we had the answers to our research questions as presented in the following Sections 5.1, 5.2, and 5.3.

5.1 On the publication of MBSE4CPS studies

In this section, we provide our results to answer RQ1 and its sub-questions.

5.1.1 Publication trends

Our answers to **RQ1.1** can be found in Figure 5, which shows how the primary MBSE4CPS studies are distributed per year. We have seen previously in Figure 2 the sharp increase of relevant papers found from the aggregated search results. However, Figure 5 gives us a closer look into the primary MBSE4CPS studies. More specifically, the primary MBSE4CPS studies were not found before 2007. The earliest primary MBSE4CPS study was found in 2007 followed by another one in 2008. Most of the primary MBSE4CPS studies were found in the last two years. 2014 recorded a peak of 13 primary studies, and 2015 saw seven primary studies published. On average, from 2007-2015, four primary studies were published annually. More recently, the period 2014-2015 saw 10 primary studies published on average annually. We agree with the opinion in [76] that model-based approaches for CPSs are gaining momentum. We can easily see a significant increase of the number of primary MBSE4CPS studies in the last two years. This increase would be a sign of the trend, in which more MBSE techniques are being developed or leveraged for the quickly expanding CPSs' popularity. Note that these numbers of publications per year are based on the official dates of publication recorded by Google Scholar, often being the dates of paper-based journal published. However, a paper that was accepted nearly the end of a year often published online first already in that year, not in the year later as officially recorded by Google Scholar. If we take a closer look at the primary studies officially published in 2014, there are two publications [63, 75] that had been published online first in 2013. Therefore, the numbers in Figure 5 should not be considered as absolute. In general, we can still see the clear increasing publication trend of the primary MBSE4CPS studies over the studied period.

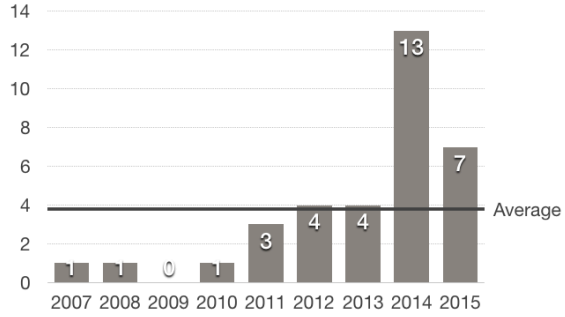

Figure 5. Distribution of primary MBSE4CPS papers per year

5.1.2 Publication venues

The bar chart in Figure 6 shows the distribution of the primary MBSE4CPS studies per venue that can give us the answers to **RQ1.2**.

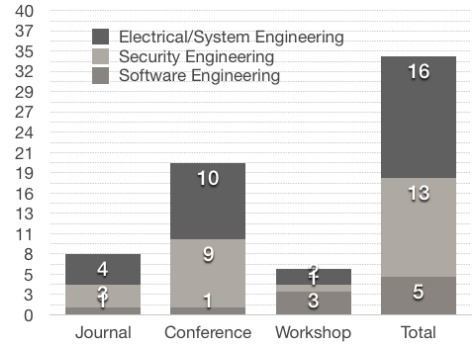

Figure 6. Distribution of papers per venue

In terms of publication venue, there are many more primary MBSE4CPS studies published at conferences (20 in total) than in journals (eight) or workshops (six). This would be understandable for a new research direction like MBSE4CPS in which ideas are supposed to be exchanged better at conferences. Besides, not many works could have been extensive or mature enough to get published in journals.

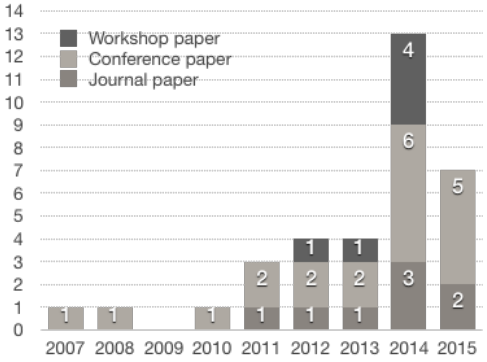

Figure 7. Distribution over publication types

Figure 7 provides a closer look at the distribution over publication types per year. The period of four years (2007-2010) contained only three conference papers of primary MBSE4CPS studies. Journal papers started to appear from 2011 and the number of journal papers seems increasing as well as the number of conference papers and workshop papers in general.

If we look at the venue types, very few primary studies (five in total) were found from software engineering related venues (see Table 3 in Appendix for more details). Publications venues that are more related to security engineering and electrical/system engineering have 13 and 16 primary MBSE4CPS studies respectively. We find that the small number of primary studies found from software engineering venues is justifiable because CPSs are relatively new research application domain for software engineering research community. The security issues for CPSs are the main focus of the primary MBSE4CPS studies whereas existing MBE techniques would only be leveraged in supporting the contributions. This could be the reason why the venues closer to security engineering and electrical/system engineering got more papers. However, we would predict that more primary MBSE4CPS studies could be found from software engineering venues in the near future when new MBE methodologies are more specifically developed for engineering the security of CPSs. It is also important to note that our classification of publication venues is not absolute as discussed later in Section 6.

5.1.3 Academia vs. industry

To answer **RQ1.3**, the pie chart in Figure 8 shows that 88 percent (30 papers) of the primary MBSE4CPS studies have authors from academia only. The shared work among academia and industry has been found in only three papers (nine percent). Only one paper [49] (three percent) is from an industrial affiliation, i.e., Roll Royce. Therefore, in total, only about 12 percent of the primary MBSE4CPS studies have the involvement from industry.

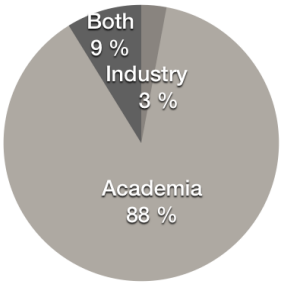

Figure 8. The academic and industrial affiliation of authors

5.1.4 Geographic distribution

For answering **RQ1.4**, we consider that a primary study was conducted in a country if the affiliation of at least an author of the primary study is in this country. For example, in a primary study that has three authors from Sweden and one author from USA, we consider that study was conducted in both Sweden and USA. Figure 9 shows that so far the researchers based in the USA (US) have involved in the biggest number of the primary MBSE4CPS studies with 12, followed by the researchers based in France (FR) with five publications, and from Canada (CA) with four publications. Researchers based in United Arab Emirates (AE), Austria (AT), Germany (DE), Sweden (SE), and United Kingdoms (UK) own three publications per country.

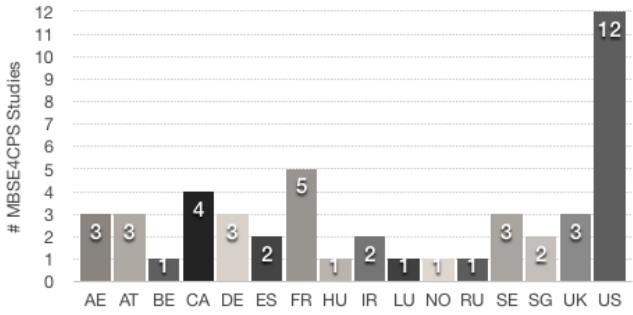

Figure 9. Number of MBSE4CPS studies per country

Researchers from Spain (ES), Iran (IR), and Singapore (SG) own two publications per country. The researchers from Belgium (BE), Hungary (HU), Luxembourg (LU), Norway (NO), Russia (RU) contributed one publication per country. The leading countries in terms of the number of primary MBSE4CPS studies such as the USA and countries in the EU are quite correlated to the research focuses on CPSs that have been being promoted in these countries and regions [25].

5.2 The characteristics of MBSE4CPS studies

This section describes the main results to answer RQ2 and its sub-questions.

5.2.1 Security concerns and security aspects

Our answers to **RQ2.1** and **RQ2.2** can be found from Figure 10. From security point of view, we would like to know how security concerns were addressed in the existing primary MBSE4CPS studies.

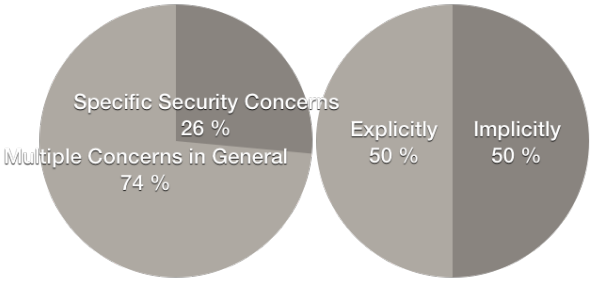

Figure 10. How security concerns were addressed in the MBSE4CPS studies

We can see in Figure 10 (left) that most approaches (74 percent, 25 papers) addressed all/multiple key security concerns (i.e., CIAAAA). This means that the security engineering activities (e.g., security analysis) are supposed to tackle all/multiple key security concerns together (either implicitly or explicitly). About a quarter (26 percent, nine papers) of the primary MBSE4CPS studies dealt with some specific security concerns, but not all the key concerns (e.g., confidentiality and integrity but not availability). Besides, Figure 10 (right) shows that half of studies did not explicitly express in their studies which specific security concerns being addressed, but rather implicitly. The reason could be simply that the authors did not explicitly mention the security concerns, or they based their approaches from security threats perspective that could be indirectly linked with security concerns. In Figure 11, we see that most of the primary MBSE4CPS studies rather focused on security analysis in general based on security threats, attacks, or vulnerabilities (76 percent). Only about 12

percent (four papers) of studies proposed solely security solutions and 12 percent proposed security solutions together with threat/attack/vulnerability analysis. More detailed analyses on these statistics are given in our answers to the other research questions as follows.

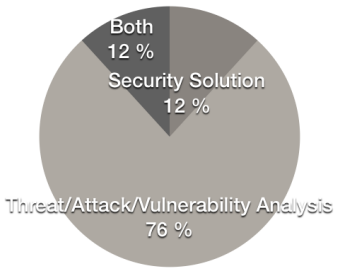

Figure 11. How security aspects were addressed in the MBSE4CPS studies

5.2.2 Modeling notation and modeling methodology

In answering **RQ2.3**, Figure 12 shows that the percentage of the primary MBSE4CPS studies that did not use the UML modeling notation (49 percent) is nearly equal to the percentage of the primary MBSE4CPS studies that used the UML modeling notation (51 percent). The modeling languages in the primary MBSE4CPS studies, which did not use the UML modeling notation, are often in forms of DSLs. Note that it is not uncommon to witness the extensive use of DSLs compared to standard modeling language, such as UML in some software engineering research area, e.g., as reported in [16]. The popularity of using DSLs in modeling (the security aspects of) CPSs that is comparable with the use of the standardized UML would reflect the heterogeneous nature of CPSs. An analysis in [43] shows that DSLs approaches for modeling CPSs could stem from various design fields such as software engineering, mechanical engineering, electrical engineering, and electronics engineering (as well as security engineering in case of MBSE4CPS). Moreover, nearly half of the primary MBSE4CPS studies leveraging non-UML modeling notations would already show the sign of a big increase in using non-UML modeling notations in security engineering. The use of non-UML modeling notations in the MDS approaches in general was only 13 percent (87 percent used the UML modeling notation) as showed in our recent relevant study [47].

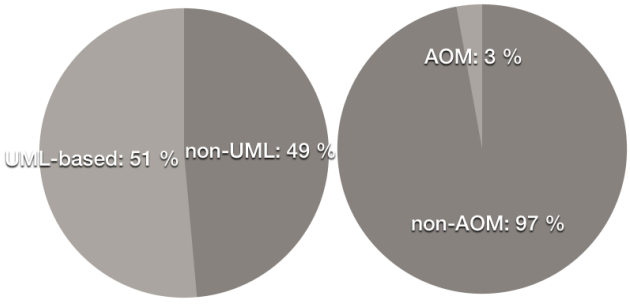

Figure 12. UML modeling notation vs. non-UML (left). AOM vs. non-AOM (right)

Table 2 classifies the primary MBSE4CPS studies according to UML-based or others. The details of the approaches using UML-based notation and not using UML-based notation are given in Table 4 and Table 5 in Appendix. Besides, we can also see that only one MBSE4CPS approach has proposed to leverage AOM. The limited use of AOM in MBSE4CPS so far is understandable as AOM was also only used in 15 percent of the primary MDS studies that were reviewed in [47].

Table 2. Primary studies classified by modeling notation

Modeling Notation/Method	Non Aspect-Oriented Modeling (non-AOM)	AOM
UML-based	[2, 4, 7, 20, 29, 35, 36, 45, 49, 51, 54, 62-65, 74, 75]	-
Others	[8, 14, 15, 19, 21, 22, 26, 27, 44, 50, 55, 58-61, 73]	[68]

Figure 13 (left) shows that the use of structural or behavioral models for specifying security threat/attack or vulnerability is less than the use of other types of models for this purpose. The other types of models are often in forms of DSMs such as attack tree model or some specific types of models that can be used by (security) analysis tools. Among the eight primary studies in total (24 percent, Figure 11) that have proposed security solutions, structural models were used more popularly (six) than behavioral models (two) and other types (three). For example, some approaches proposed security patterns for CPSs that normally expressed in structural models. However, the number of models used for specifying security solutions is much smaller than the number of models for specifying threats/attacks and vulnerabilities. The reason is that only 24 percent of the primary studies proposed security solutions compared to 76 percent of the primary studies proposed threat/attack/vulnerability analysis only (Figure 11).

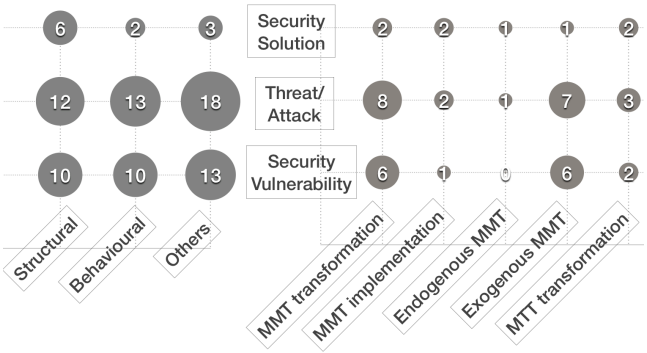

Figure 13. Types of models and model transformations

Figure 13 (right) shows that MMTs were leveraged in a few primary studies, i.e., addressing threat/attack (eight primary studies), vulnerability (six), and security solution (two). In total, the number of primary studies that mentioned to leverage MMTs is eight, and the total number of primary studies that mentioned to have MTTs is three. These numbers are quite small compared to 34 papers of the primary MBSE4CPS studies. Moreover, only two primary studies (proposing security solutions together with threat/attack analysis) provided some implementation information of MMTs [19, 55]. Also, we can see in Figure 13 (right) that out of eight primary studies having leveraged MMTs, most of them (seven) are of type exogenous MMTs to transform security/system models to some other DSMs that can be used by (security) analysis tools (e.g., [19, 51]). Only one endogenous MMT was used to compose security aspects into system model [55]. There is no approach that provided automated security analyses directly on security/system models (at verification stage) because (model-based) formal (security) analyses would require specialized analysis methods with specialized model inputs. Therefore, exogenous MMTs have been developed in a few primary studies

to bridge the gap. In other words, exogenous MMTs could help for transforming security/system models into specialized models that are closer to the inputs of verification/analysis methods and/or tools.

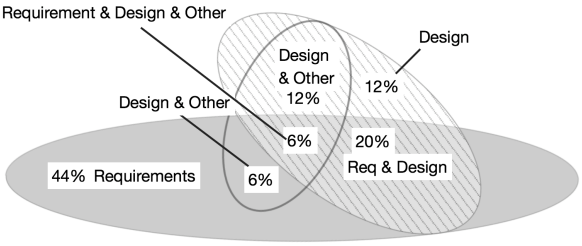

Figure 14. The distribution of MBSE4CPS studies regarding the main stages of the security development lifecycle

5.2.3 Security engineering phases and tool support

The focus of the primary MBSE4CPS studies on security analysis can also be explained when we look closer to know which stage(s) of the SDL that the primary MBSE4CPS studies worked on. In answering **RQ2.4,** Figure 14 shows the distribution of the primary MBSE4CPS studies according to the main stages: requirements/domain analysis, architecture/design, and others (implementation, verification, release, response). 100 percent of the primary MBSE4CPS studies worked on either the requirements/domain analysis or architecture/design or both stages. Nearly half of the studies (44 percent) focused solely at the requirements stage. Only six percent proposed more complete security development approaches from requirements/domain analysis to architecture/design, and then to later stages. More detailed analyses of the MBSE4CPS studies according to the main stages of the SDL are provided later with Figure 18.

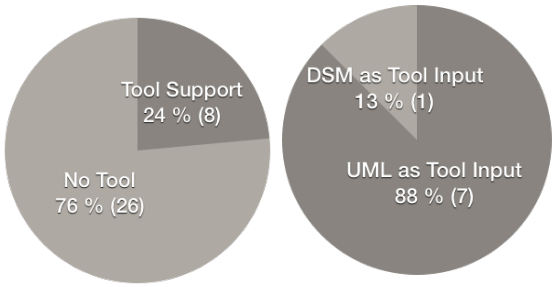

Figure 15. Tool support statistics and tool inputs

In terms of tools support (e.g., for modeling, security analysis), eight primary studies (24 percent) have mentioned tools support, in which only two proposes a new tool and its extended version [29, 65]. Six primary studies are based on extending existing tools. 26 primary studies (76 percent) do not propose any tool support (Figure 15).

Tool platform. Considering only the tools, there is a common combination of UML-based modeling tools with analysis tools for building a tool platform in five of them (62.5 percent). More precisely, UML-based modeling tools such as Papyrus in [7], MagicDraw in [63] are in combination with (formal) analysis tools such as ProVerif in [51], Diversity in [7]. The rest of tools are not described clearly. The percentage of papers not using tool support could be quite high for the field of study. Especially in terms of security engineering, tool support is crucial. For example, tools are needed for security engineers to better use security-oriented DSLs for specifying security models. Unlike UML-based modeling that has a range of available tool supports (e.g., Papyrus[8], MagicDraw[9]), DSLs are often tailored and developed from scratch. Without tool support, DSLs' end-users would not be able to use DSLs properly for specifying security models. Tools are also vital to support automated security verification and validation. Doing security verification and validation manually could be very error-prone because security implementations are often scattered and tangled throughout system.

Tool Input. The total number of primary studies using the UML-based notation as input is seven (88 percent of the papers reporting a tool, Figure 15). Only one primary study with tool support (i.e., [19]) uses non-UML-based DSMs as tool input. The number of UML-based approaches with tool support is dominant compared to the non-UML ones with tool support. This is understandable because tool support for UML has been matured and industrialized.

Tool Output. Tool outputs are in the forms of security analyses results such as security proofs (one tool), security risks (vulnerabilities, six), and security requirements based on risk estimation (one). The forms of tool outputs are matched with the observation that the majority of the primary MBSE4CPS studies are mainly for threat, attack or vulnerability analysis (Figure 11).

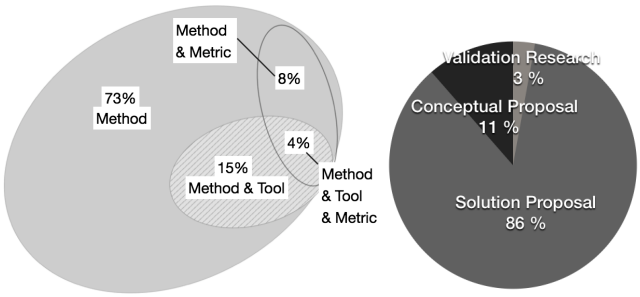

Figure 16. The types of MBSE4CPS research contribution (left) and the fine-grained types of research (right)

5.2.4 Research contribution and research type

Figure 16 and Figure 17 can help us to answer **RQ2.5**. Figure 16 (left) shows that method (e.g., a security analysis method) is the main type of research contribution in all the primary MBSE4CPS studies. Among the primary studies, 73 percent introduced solely methods, 15 percent introduced methods together with tool support, eight percent introduced methods together with (security) metric(s), and four percent introduced methods, tool, and metrics in the same study. We do not show in Figure 16 (left) the percentage of the primary MBSE4CPS studies that have models as part of research contribution because it is obvious from our selection criteria that all the primary MBSE4CPS studies must have model(s) in their research contributions.

In terms of fine-grained types of research, Figure 16 (right) shows that 86 percent of the primary MBSE4CPS studies are of the type of solution proposals whereas only three percent (one paper) is of the type of validation research [59]. 11 percent of studies are of the type of conceptual proposals only. None of the type of opinion, evaluation study or experience report was found.

[8] Papyrus Modeling Environment, https://eclipse.org/papyrus/

[9] MagicDraw Modeling Tool, http://www.nomagic.com/products/magicdraw.html

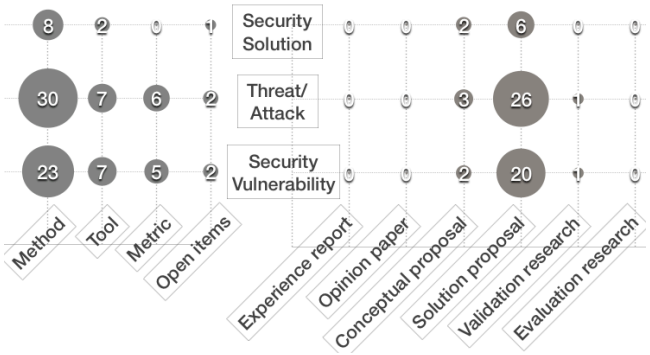

Figure 17. Security aspects w.r.t. research contributions (left) and types of research (right)

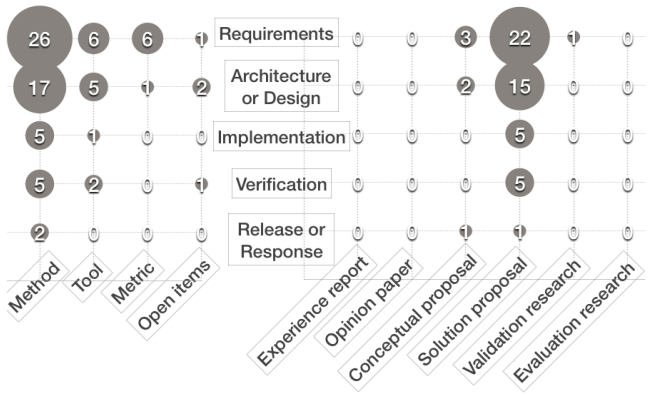

Figure 18. Engineering stages w.r.t. research contributions (left) and types of research (right)

To have more detailed information, we analyze the distribution of papers among various analysis aspects. Figure 17 and Figure 18 show that most of the primary MBSE4CPS studies have research contribution type as methods (e.g., for security analysis) related to threat/attack or vulnerability (30 and 23 papers respectively, Figure 17 left) with a few tools introduced (seven and seven), and at early stages such as requirement/domain analysis, or design (26 and 17 papers, Figure 18 left), also with a few same tools (six and five). Only about a quarter of the primary MBSE4CPS studies (eight papers, Figure 17 left) have research contribution as method for security solution and only two have tool support. Similarly, Figure 18 (left) shows that very few methods supported the later stages of SDL such as implementation (five), verification (five), and release (two), with very limited tool support (only one or two). Verification and validation stage is very important for providing evidence and evaluation on the security of CPSs. Because the security of CPSs is critical as discussed before, the security verification and validation stage must be a vital part of the SDL. However only a few approaches (e.g., [19, 51]) proposed formal verification, and not any studies proposed model-based security testing approach for CPSs. There are two papers raised some open items/issues, i.e., how security analysis can be integrated with different classes of DSLs (for specifying threat/attack/vulnerability) such as those based on control flow [19], or the challenges in bridging the implementation gap from requirements to design and then to real (hardware based) implementations for the security of CPSs at the code level [4].

On the right hand side of Figure 17 and Figure 18, we can see that most of the primary MBSE4CPS studies have research type as solution proposal (e.g., for analysis) related to threat/attack or vulnerability (26 and 20 papers respectively, Figure 17), and at early stages such as requirement/domain analysis (22 papers, Figure 18) and design (15). Also less than a quarter of the primary MBSE4CPS studies (six, Figure 17) have research type as solution proposal containing security solution, and at the later stages of SDL (five, Figure 18). Only one paper of type validation research has been found [59], which was mainly about assessing an approach for security requirements engineering via an academic case study of smart grid (considering threat/attack/vulnerability). We have not found any evaluation research or experience report, or opinion paper.

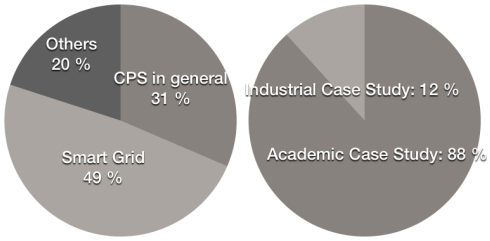

Figure 19. Application domains and case studies

5.2.5 Application domains and uncertainty

In answering **RQ2.6**, the pie chart on the left of Figure 19 shows that nearly half of the primary MBSE4CPS studies (49 percent) used the smart energy grids as case studies or application domains. This is understandable because smart grid (the next-generation power system) could be the most popular instance of CPSs that is receiving national priorities in many developed countries such as the USA, and in the EU [25]. A recent survey also shows that most intrusion detection techniques for CPSs have been proposed so far are for the security of smart utility (mainly smart energy grid) [42]. Moreover, to realize many of its advanced features, smart grid depends heavily on (open) information networking that inevitably makes it more vulnerable to security threats [67]. In smart grids, information and communication technology (ICT) is increasingly integrated throughout the grid to support novel communication and control functions but at the same time bring up lots of ICT security challenges. About one third of the primary MBSE4CPS studies (31 percent) are for CPSs in general, e.g., a generic language for describing attacks on CPSs [74]. Other application domains of the primary MBSE4CPS studies accounted for 20 percent were varied including automotive CPSs, healthcare, and transportation. Figure 19 (right) shows that most of the primary MBSE4CPS studies (88 percent) were only evaluated on academic case studies (e.g., many academic case studies are smart grids) whereas a much smaller number of primary studies (12 percent) had industrial case studies.

To answer **RQ2.7**, while reviewing the primary MBSE4CPS studies we also paid attention to check if any primary MBSE4CPS study dealt with the security of CPSs taking into account uncertainty. However, we did not find any paper addressing the uncertainty problem of CPSs.

5.3 Open issues & proposed research agenda

Based on our findings for the research questions **RQ1** and **RQ2**, we would like to point out the current open issues of MBSE4CPS research for answering **RQ3.1** and **RQ3.2**. For each open issue, we propose some (research) directions to address the issue.

5.3.1 Implicit security concerns/objectives

In Section 5.2, our answer to RQ2.1 states that half of existing the primary MBSE4CPS studies did not explicitly express in their studies what specific security concerns (e.g., CIA) being

addressed, but rather implicitly. From security engineering point of view, security approaches must be driven by security concerns. By explicitly pointing out what security concerns are being addressed, the primary MBSE4CPS studies could deal with those concerns more systematically and convincingly. Therefore, we would suggest that the security concerns to be referenced explicitly in every primary MBSE4CPS study. It could be that currently a common understanding of security and CPSs together is missing. One way to achieve this is to develop a conceptual model that can cover both the aspects together.

5.3.2 Very few security solutions engineered

As pointed out in our answer for RQ2.2, most of the primary studies focused on supporting security analyses based on security threats, attacks, or vulnerabilities and did not focus on engineering security solutions. It can be understood that as a relatively new field, MBSE4CPS research so far focused mainly on requirements and domain analysis. Therefore, security solutions for CPSs are still rare. In addition, there could be new types of security threats, attacks that are very different from traditional ones in many CPSs' application domains, e.g., new security threats to the physical parts of CPSs. The security solutions for these new kinds of threats would still be under development. In the future, more new MBSE4CPS approaches should be proposed for engineering security solutions in the development of CPSs.

5.3.3 Limited automation in formal security analysis

As discussed in Section 5.2, there was no primary MBSE4CPS study that supports analyses directly on security/system models at the verification stage. Some (e.g., [19, 51]) discussed about translating models into other formalisms for enabling automated analyses. Transformation into other formalisms for analyses poses additional overhead of translation that may not be fully possible and may not be fully automated. However, this transformation approach provides access to mature analyses tools, such as based on Alloy [41]. The employment of model transformations in the primary MBSE4CPS studies was very limited and could be leveraged more. Nevertheless, model transformations could be considered as the heart and soul of model-driven software development in general [56]. Model transformations would have been used more extensively, e.g., for enabling automated analyses than they are currently used in a few primary MBSE4CPS studies. Based on our findings, we believe that the current MBSE4CPS literature is immature in terms of providing automated formal analyses at the verification stage. This limitation can also be seen in terms of the very limited tool support at this stage proposed by the existing primary MBSE4CPS studies. Also discussed in the results, among a few primary studies that propose tools support, it is quite common that UML-based modeling tools are combined with (formal) analysis/verification tools. The combination of DSL-based modeling tools with analysis tools was very rare even among a few primary MBSE4CPS studies with tool support.

5.3.4 Limited work on the later stages of SDL

Since the area of security research in CPSs is still very immature, most of the primary studies focused on analyses in the early stages of SDL (i.e., requirement, design) as discussed in Section 5.2. We believe that as the field matures, we expect to see more support for security engineering in the later stages of SDL such as implementation, verification (e.g., model-based security testing, model-based formal verification), release, and response. Verification stage is very important for verifying the security of CPSs. In our answers to RQ2.5, we pointed out that only a few primary studies proposed formal verification, and not any primary studies proposed model-based security testing (MBST) approach for CPSs. MBST would be a potential direction to contribute for the validation of the security of CPSs.

5.3.5 Limited work on validation, evaluation studies

As discussed in Section 5.2, we could not find any primary study of type opinion, evaluation study or experience report. Most of the primary MBSE4CPS studies are solution proposals whereas only one paper is of the type of validation research but more for requirements engineering than MBSE4CPS [59]. Once again, this gives a clear indication that the MBSE4CPS field is immature.

5.3.6 Limited collaboration with industry

As showed in Section 5.1 and Section 5.2, most of the primary MBSE4CPS studies were only evaluated on the academic case studies whereas a much smaller number of the primary studies (about ten percent) were based on real industrial case studies. Considering the trend that CPSs are driving the so-called fourth industrial revolution, evaluations on the real industrial case studies should be seriously needed. Besides, most of the primary studies have authors from academia only, which would imply the lack of collaboration in MBSE4CPS research between academia and industry. Therefore, more collaboration among academia and industry for MBSE4CPS research needs to be promoted.

5.3.7 The lack of dealing with uncertainty

Uncertainty is inherent in CPSs due to CPSs' complexity and multidisciplinary nature, e.g., in the integration of different technologies in computing, networking, and control to monitor and control not only information but also physical processes [30]. In addition, security issues in the context of CPSs could be one of the key contributors to introducing uncertainty in CPSs that may lead to their unreliable or even unsafe operations. The tight interaction between cyber and physical parts of CPSs as well as the heavy dependence on (more open) communication network make CPSs, especially its physical processes, more vulnerable to the security vulnerabilities in the cyber side [66]. On the other hand, inadequate security constraints (e.g., access control) may fail some physical processes that have critical real time requirement. Uncertainty is not handled in general in the context of CPSs and consequently uncertainty due to security related issues has not been studied at all as it is demonstrated by our SMS. MBSE4CPS research community should spend more effort to tackle uncertainty problems for CPSs, especially for the security of these important systems.

5.3.8 Modeling and integration challenges

Nearly half of the primary MBSE4CPS studies leveraging non-UML modeling notations would already show the trend of using domain-specific languages in engineering (the security of) CPSs. Modeling a CPS itself is challenging due to its multi-disciplinary nature requiring expertise in software, hardware, and physical phenomena to name a few. (Non UML-based) DSLs are worth to be explored in the MBSE4CPS studies because each DSL is normally lightweight (compared to general modeling languages) and tailored for engineering a specific problem domain in software, or hardware, or security of CPSs. Developing and combining DSLs could be a promising solution for the MBSE4CPS studies to tackle the multi-disciplinary nature in engineering CPSs and security. Besides, the development of UML profiles as DSLs is also a possibility for the approaches that are based on the UML modeling notation as surveyed in [47]. In fact, some of the primary MBSE4CPS studies (e.g., [4, 36]) have proposed to extend UML-based system modeling languages such as SysML and MARTE.

Another open challenge would be the integration of different classes of DSLs (for specifying security aspects) with security analysis (also pointed out in [19]). Model transformations could help bridging this gap but will need to be investigated more in this context. Combining modeling and analyses of security concerns together with CPSs is even more challenging. In most cases, security concerns are crosscutting concerns that pose additional modeling challenges. A promising modeling paradigm to address this challenge is AOM. So far only one primary MBSE4CPS study [68] proposed to leverage AOM and this direction is indeed very open.

6. THREATS TO VALIDITY

It is essential to have explicit discussion of the limitations of a SLR itself besides presenting its results [18]. Even though a SMS would have less in-depth analysis than a SLR, we still discuss some threats to validity of our study as follows.

There are different kinds of support for using keywords in searching for papers in different online databases. We had to adapt the use of search terms according to different search functionalities and search refinement processes provided by different online databases. We tried to complement the limitations of database search by conducting an extensive snowballing process as presented in Section 3.4.2.

We are aware that some systems classified under different categories such as SoS, embedded systems, distributed systems that could implicitly be CPSs. We had to check carefully the case studies in many candidate papers (e.g., [7] and [64]) to see if they are some kinds of CPSs. For each candidate paper facing this classification challenge, we had discussion among the authors to reach an inclusion/exclusion decision.

As discussed in [52], it is difficult to be consistent in classifying research types with the research types proposed from [70]. We used the decision table in [52] to disambiguate the classification of studies.

Many publication venues could have papers from different related research domains such as software engineering, security engineering, and electrical engineering. Therefore, many venues would belong to multiple domains but we classified a venue to the closest research domain based on the description of the venue, the relevant calls for papers submission, and our subjective opinions. Therefore, our classification of publication venues is not absolute.

The set of primary MBSE4CPS studies could not be very big to have more generalized results but we would suppose that by analyzing this set, we could have shed some light into an emerging, important, and challenging research area such as MBSE4CPS.

7. RELATED WORK

In [46] and [47], the model-driven development of secure systems in general, not specifically for CPSs, was extensively reviewed. The focus was model-driven development, not in a broader scope as model-based engineering. In other words, these studies examined the Model-Driven Security approaches (for all application domains) classified as Model-Driven Development in [12], in which models "drive" development process. This SMS study examined the MBSE approaches (for CPSs only) classified as MBE in [12], in which models could be engineered at any single stage in the development life cycle, and do not necessarily drive the development process. There is one primary study (i.e., [19]), which is common among this SMS, [46], and [47].

Model-based techniques for systems of systems (SoS) engineering were surveyed in [48]. More specifically, the authors examined the model-based techniques for SoS description, simulation, testing, and verification. The focus of [48] was SoS which would have a bigger scope than CPSs. Some CPSs could be in a subset of SoS. Besides, [48] did not specifically address the security of SoS. Moreover, the papers surveyed in [48] were selected solely based on the personal awareness of the authors, not via a systematic search and selection process as in our SMS. There is not any primary MBSE4CPS study surveyed in [48].

In [76], the authors assessed the state of the art and the state of the practice in the verification and validation of CPSs. Their study methodology is twofold: a literature review of CPSs' verification and validation; and a structured on-line survey plus semi-structured interviews. MBE for the verification and validation of CPSs is one of the categories in their literature review part. Their study is not about the security of CPSs. Also, there is not any primary MBSE4CPS study discussed in [76].

Testing approaches that are specific for CPSs have been surveyed recently in [5]. A few model-based testing approaches for CPSs were discussed. However, none of the testing approaches in the survey addresses the security of CPSs.

8. CONCLUSION AND FUTURE WORK

8.1 Conclusion

In this paper, we have presented the results of a systematic mapping study on the existing model-based security engineering studies for cyber-physical systems (MBSE4CPS). The results could shed some light on an emerging research area, which is interdisciplinary among research domains such as system engineering, software engineering, and security engineering. More specifically, our study was designed and conducted based on a rigorous SMS protocol for identifying a set of primary MBSE4CPS studies to answer three big research questions. The main contributions of this paper are our answers to these questions and their sub-questions, which are summarized as follows:

RQ1. What are the publication statistics of the existing primary MBSE4CPS studies in the literature?
(In answering *RQ1.1*) The first primary MBSE4CPS study was published in 2007. On average, from 2007-2015, four primary studies were published annually. The number of the primary MBSE4CPS studies has significantly increased (ten on average) during the last two years (2014-2015), which could mean this research area is expanding. (*RQ1.2*) In terms of publication venue, there are more primary MBSE4CPS studies published at conferences than in journals or workshops. Fewer primary studies were found from software engineering related venues compared to security engineering and system engineering. (*RQ1.3*) Most of the primary MBSE4CPS studies have authors from academia only. The involvement of industry has been found in very few primary studies. (*RQ1.4*) So far the researchers based in the USA have involved in the most primary MBSE4CPS studies, followed by the researchers based in France, Canada, and 13 other countries mainly in Europe. The leading countries in terms of the number of the MBSE4CPS primary studies such as the USA and countries in the EU are quite correlated to the research focuses on CPSs that have been being promoted in these countries.

RQ2. What are the existing primary MBSE4CPS studies and their characteristics?
(RQ2.1) Most of the primary studies addressed multiple key security concerns. However, half of the primary studies did not

explicitly express in their studies which specific security concerns being addressed, but rather implicitly. *(RQ2.2)* In fact, most of the primary studies focused on security analysis in general based on security threats, attacks, or vulnerabilities. Only about one tenth of the primary studies proposed solely security solutions and one-tenth proposed security solutions together with threat/attack/vulnerability analysis. *(RQ2.3)* The use of domain-specific languages (DSLs) in the primary MBSE4CPS studies is comparable with the use of the standardized UML. The use of structural or behavioral models for specifying security threat/attack or vulnerability is less than the use of other types of models (e.g. created in DSLs) for this purpose. The number of models used for specifying security solutions is much smaller than the number of models for specifying threats/attacks and vulnerabilities. Model-to-model transformations (MMTs) were leveraged in quite small number of the primary MBSE4CPS studies. Fewer provided some implementation information of MMTs. *(RQ2.4)* As an emerging field, MBSE4CPS research so far focused on the early stages of the security development lifecycle (SDL) such as requirement engineering and analysis. All the primary MBSE4CPS studies worked on either the requirements/domain analysis or architecture/design or both stages. Nearly half of the primary studies focused solely on requirements stage. Very few proposed more complete security development approaches from requirements/domain analysis to architecture/design, and then to later stages. In terms of tools support, less than one-third of the primary MBSE4CPS studies have mentioned tools support.
(RQ2.5) Method (e.g., a security analysis method) is the main type of research contribution in all the primary MBSE4CPS studies. Among the primary studies, most introduced solely method. Few introduced methods together with tool support, or metric(s). Fewer introduced method, tool, and metric in the same study. Most of the primary studies are of type research solution proposal whereas only one is of type validation research. About one-tenth of the primary studies are of type conceptual proposal only, and none of type opinion, evaluation study or experience report was found. Very few methods supported the later stages of SDL such as implementation, verification, and release. *(RQ2.6)* Nearly half of the primary MBSE4CPS studies used the smart energy grids as case studies or application domains. About one third of the primary MBSE4CPS studies are for CPSs in general, e.g., a generic language for describing attacks on CPSs. Other application domains of the primary MBSE4CPS studies accounted for one-fifth were varied including automotive CPSs, healthcare, and transportation. Most of the primary MBSE4CPS studies were only evaluated on academic case studies whereas much smaller number of primary studies had industrial case studies. *(RQ2.7)* We kept in mind to check if any primary study has addressed the uncertainty aspect of CPSs but did not find any.

RQ3. What are the open issues of MBSE4CPS research?
First, half of the existing primary MBSE4CPS studies did not explicitly express in their studies what specific security concerns (e.g., CIA) being addressed. It could be that currently a common understanding of security and CPSs together is missing. One way to achieve this is to develop a conceptual model that can cover the both aspects together. Second, most of the primary studies focused on supporting security analyses based on security threats, attacks, or vulnerabilities and did not focus on engineering security solutions. More MBSE4CPS studies should be proposed with security solutions in the later stages of the SDL such as implementation and verification. Third, not any primary MBSE4CPS study that supports analyses directly on the security/system models at verification stage. The current MBSE4CPS literature is immature in terms of providing automated formal analyses at verification stage. This limitation can also be seen in terms of very limited tool support proposed by the existing primary MBSE4CPS studies. Fourth, we also found that the collaboration between academia and industry as well as the involvement of industry in this research area so far is very limited. Besides, the lack of dealing with uncertainty is worth to note because uncertainty would be inevitable in real CPSs and tangle with their security. Fifth, modeling CPSs itself is challenging due to its multi-disciplinary nature. DSLs could be a key part in engineering (the security of) CPSs in their multidisciplinary nature. However, an open challenge would be the integration of different DSLs, e.g., by leveraging model transformations.

8.2 Future work

Our SMS protocol and the set of primary MBSE4CPS studies could be used in a follow-up SMS that reports more up-to-date results based on the primary MBSE4CPS studies reported in this paper plus newly found primary MBSE4CPS papers in the future. The set of primary MBSE4CPS papers could be enriched and updated in three ways. First, new primary MBSE4CPS studies could be found from new database searches that cover the period after this SMS, i.e. from 2016 on. Second, we would expect many more MBSE4CPS studies in the future as well as more specific or dedicated publication venues for publishing MBSE4CPS studies. If so, one could conduct manual search on those venues to find new primary MBSE4CPS studies. Third, one could conduct another recursive snowballing, especially forward snowballing (by checking citations), on the set of primary MBSE4CPS studies including newly found ones. After the set of primary MBSE4CPS studies is updated, our protocol can be reused and adopted to extract, synthesis data, and report on the updated results.

On the other hand, we plan to do a systematic review more deeply into the model-based security verification and validation approaches for CPSs (MBSVV4CPS), a follow up from this SMS. The set of primary MBSE4CPS papers can be updated as discussed above, and all MBSVV4CPS studies (a subset of MBSE4CPS) can be filtered out and reviewed in details. Besides, because the uncertainty aspects of CPSs have not been tackled, we are developing a model-based security testing approach for CPSs that takes into account uncertainty.

ACKNOWLEDGMENTS

This research was supported by RCN funded MBT4CPS project. Phu-Hong Nguyen, Tao Yue, and Shaukat Ali are also supported by the EU Horizon 2020 funded project U-Test (Testing Cyber-Physical Systems under Uncertainty). Tao Yue and Shaukat Ali are also supported by RCN funded Zen-Configurator project, RFF Hovedstaden funded MBE-CR project, and RCN funded Certus SFI.

APPENDIX

The following tables present extra information about the publication venues of the primary MBSE4CPS studies, and the list of these primary studies. Table 3 shows all those publication venues with their types (journal, conference, workshop), their domains (software engineering, system engineering, security engineering), and the number of primary studies found from there.

Table 3. Distribution of publication venues

Publication venue	Type	Domain	# MBSE4CPS

			papers
Information Systems	Journal	System Engineering	1
Critical Infrastructure Protection	Journal	Security Engineering	1
Computer Science - Research and Development	Journal	System Engineering	1
Systems Engineering	Journal	System Engineering	1
Security and Communication Networks	Journal	Security Engineering	1
Requirement Engineering	Journal	Software Engineering	1
Advances in Security	Journal	Security Engineering	1
IEEE Transactions on Smart Grid	Journal	System Engineering	1
International Conference on Computer Safety, Reliability and Security (SAFECOMP)	Conference	Security Engineering	1
International Conference on Model-Driven Engineering and Software Development (MODELSWARD)	Conference	Software Engineering	1
International Conference on Information Systems Security and Privacy (ICISSP)	Conference	Security Engineering	1
IEEE Pacific Rim International Symposium on Dependable Computing (PRDC)	Conference	Security Engineering	1
IEEE European Modeling Symposium (EMS)	Conference	System Engineering	1
Sensors and Systems for Space Applications	Conference	System Engineering	1
International Symposium for Industrial Control System & SCADA Cyber Security Research (ICS-CSR)	Conference	System Engineering	2
International ISC Conference on Information Security and Cryptology	Conference	Security Engineering	2
HASE: High-Assurance Systems Engineering	Conference	Security Engineering	1
Design Automation Conference (DAC)	Conference	System Engineering	1
Asia-Pacific Council on Systems Engineering (APCOSEC)	Conference	System Engineering	1
International Conference on Parallel, Distributed and Network-based Processing (PDP)	Conference	System Engineering	1
International Symposium on Resilient Control Systems (ISRCS)	Conference	Security Engineering	1
International Conference on Secure Software Integration and Reliability Improvement - Companion (SSIRI-C)	Conference	Security Engineering	1
New Technology Distribution (NOTERE)	Conference	System Engineering	1
Hawaii International Conference on System Sciences (HICSS)	Conference	System Engineering	1
IEEE Power and Energy Society (PES)	Conference	System Engineering	1
Engineering of Computer Based Systems (ECBS)	Conference	System Engineering	1
International Conference on Software Testing Workshops (ICSTW)	Workshop	Software Engineering	1
Smart Grid Security (SmartGridSec)	Workshop	Software Engineering	2
International Conference on Quantitative Evaluation of SysTems - Workshops (QEST-W)	Workshop	System Engineering	1
Cyber Security and Information Intelligence Research Workshop (CSIIRW)	Workshop	Security Engineering	1
International Conference on Dependable Systems and Networks - Workshops (DSN)	Workshop	System Engineering	1

Table 4 lists all the primary studies, in which UML-based modeling notation is the primary notation used.

Table 4. Papers using UML-based notation

Author(s)	Title	Research type	Contribution Type
Oates, Thom [49]	*Security-aware, model-based systems engineering with SysML*	Conceptual Proposal	Method
Zafar, Arnautovic [75]	*System security requirements analysis: A smart grid case study*	Solution Proposal	Method
Vasilevskaya, Gunawan [63]	*Integrating security mechanisms into embedded systems by domain-specific modelling*	Solution Proposal	Method, Tool
Knirsch, Engel [35]	*Privacy Assessment of Data Flow*	Solution Proposal	Method, Metric

	Graphs for an Advanced Recommender System in the Smart Grid		
Abdallah, Motii [2]	*Using Model Driven Engineering to Support Multi-paradigms Security Analysis*, in *Model-Driven Engineering and Software Development*	Solution Proposal	Method
Jauhar, Binbin [29]	*Model-Based Cybersecurity Assessment with NESCOR Smart Grid Failure Scenarios*	Solution Proposal	Method, Tool, Metric
Ur-Rehman and Zivic [62]	*Secure Design Patterns for Security in Smart Metering Systems*	Solution Proposal	Method
Vasilevskaya and Nadjm-Tehrani [64]	*Quantifying Risks to Data Assets Using Formal Metrics in Embedded System Design*	Solution Proposal	Method, Metric
Vu, Tippenhauer [65]	*CyberSAGE: a tool for automatic security assessment of cyber-physical systems*	Solution Proposal	Method, Tool
Yampolskiy, Horváth [74]	*A language for describing attacks on cyber-physical systems*	Conceptual Proposal	Method
Bannour, Escobedo [7]	*Designing Sequence Diagram Models for Robustness to Attacks*	Solution Proposal	Method, Tool
Neureiter, Eibl [45]	*A concept for engineering smart grid security requirements based on SGAM models*	Conceptual Proposal	Method, Tool, Metric
Lemaire, Lapon [36]	*A SysML extension for security analysis of industrial control systems*	Solution Proposal	Method
Apvrille and Roudier [4]	*SysML-Sec: a SysML environment for the design and development of secure embedded systems*	Solution Proposal	Method, Tool, Open Issue
Ruiz, Harjani [54]	*A methodology for the analysis and modeling of security threats and attacks for systems of embedded components*	Solution Proposal	Method
Pedroza, Apvrille [51]	*Avatar: A sysml environment for the formal verification of safety and security properties*	Solution Proposal	Method, Tool
Fernandez and Larrondo-Petrie [20]	*Designing secure SCADA systems using security patterns*	Conceptual Proposal	Method

Table 5 lists all the primary studies, in which non-UML-based modeling notations are mainly used.

Table 5. Papers using non-UML notation

Author(s)	Title	Research type	Contribution Type
Forbes, Vu [22]	*Defending a nanosatellite cyber-physical system*	Solution Proposal	Method
Nasr and Varjani [44]	*Petri net model of insider attacks in SCADA system*	Solution Proposal	Method
Orojloo and Azgomi [50]	*A method for modeling and evaluation of the security of cyber-physical systems*	Solution Proposal	Method, Metric
Tabrizi and Pattabiraman [61]	*A model-based intrusion detection system for smart meters*	Solution Proposal	Method
Beckers, Faßbender [8]	*A threat analysis methodology for smart home scenarios*, in *Smart Grid Security*	Solution Proposal	Method
Hartmann, Fouquet [27]	*Reactive security for smart grids using models@ run. time-based simulation and reasoning*	Solution Proposal	Method
Hahn and Govindarasu [26]	*Model-based intrustion detection for the smart grid (MINDS)*	Solution Proposal	Method
Saadatmand, Cicchetti [55]	*Managing Timing Implications of Security Aspects in Model-Driven Development of Real-Time Embedded Systems*	Solution Proposal	Method
Tabrizi and Pattabiraman [60]	*A model for security analysis of smart meters*	Solution Proposal	Method
Yampolskiy, Horvath [73]	*Systematic analysis of cyber-attacks on CPS-evaluating applicability of DFD-based approach*	Solution Proposal	Method
Chen, Sanchez-Aarnoutse [14]	*Petri net modeling of cyber-physical attacks on smart grid*	Solution Proposal	Method
Fletcher and Liu [21]	*Security requirements analysis, specification, prioritization and policy development in cyber-physical systems*	Solution Proposal	Method, Metric

Cheung, Hamlyn [15]	*Role-based model security access control for smart power-grids computer networks*	Solution Proposal	Method
Eby, Werner [19]	*Integrating security modeling into embedded system design*	Solution Proposal	Method, Tool, Metric
Suleiman, Alqassem [58]	*Integrated smart grid systems security threat model*	Solution Proposal	Method
Suleiman and Svetinovic [59]	*Evaluating the effectiveness of the security quality requirements engineering (SQUARE) method: a case study using smart grid advanced metering infrastructure*	Validation Research	Method

REFERENCES

[1] (NSF), N.S.F. *Cyber-Physical Systems (CPS) PROGRAM SOLICITATION NSF 16-549*. 2016 [cited 2016 May 2016]; Available from: http://www.nsf.gov/pubs/2016/nsf16549/nsf16549.htm.

[2] Abdallah, R., A. Motii, N. Yakymets, and A. Lanusse, *Using Model Driven Engineering to Support Multi-paradigms Security Analysis*, in *Model-Driven Engineering and Software Development*. 2015, Springer. p. 278-292.

[3] Ali, S. and T. Yue. *U-Test: Evolving, Modelling and Testing Realistic Uncertain Behaviours of Cyber-Physical Systems*. in *Software Testing, Verification and Validation (ICST), 2015 IEEE 8th International Conference on*. 2015. IEEE.

[4] Apvrille, L. and Y. Roudier, *SysML-Sec: a SysML environment for the design and development of secure embedded systems*. APCOSEC 2013, 2013.

[5] Asadollah, S.A., R. Inam, and H. Hansson, *A Survey on Testing for Cyber Physical System*, in *Testing Software and Systems*. 2015, Springer. p. 194-207.

[6] Balaji, B., A. Faruque, M. Abdullah, N. Dutt, R. Gupta, and Y. Agarwal. *Models, abstractions, and architectures: the missing links in cyber-physical systems*. in *Proceedings of the 52nd Annual Design Automation Conference*. 2015. ACM.

[7] Bannour, B., J. Escobedo, C. Gaston, P. Le Gall, and G. Pedroza. *Designing Sequence Diagram Models for Robustness to Attacks*. in *Software Testing, Verification and Validation Workshops (ICSTW), 2014 IEEE Seventh International Conference on*. 2014. IEEE.

[8] Beckers, K., S. Faßbender, M. Heisel, and S. Suppan, *A threat analysis methodology for smart home scenarios*, in *Smart Grid Security*. 2014, Springer. p. 94-124.

[9] Bézivin, J., *Model driven engineering: An emerging technical space*, in *Generative and transformational techniques in software engineering*. 2006, Springer. p. 36-64.

[10] Biolchini, J., P.G. Mian, A.C.C. Natali, and G.H. Travassos, *Systematic review in software engineering.* System Engineering and Computer Science Department COPPE/UFRJ, Technical Report ES 679.05, 2005.

[11] Bishop, M., *Computer security: art and science*. Vol. 200. 2012: Addison-Wesley.

[12] Brambilla, M., J. Cabot, and M. Wimmer, *Model-Driven Software Engineering in Practice*. 2012: Morgan & Claypool Publishers.

[13] Cardenas, A.A., S. Amin, and S. Sastry. *Secure control: Towards survivable cyber-physical systems*. in *The 28th International Conference on Distributed Computing Systems Workshops*. 2008. IEEE.

[14] Chen, T.M., J.C. Sanchez-Aarnoutse, and J. Buford, *Petri net modeling of cyber-physical attacks on smart grid.* IEEE Transactions on Smart Grid, 2011. **2**(4): p. 741-749.

[15] Cheung, H., A. Hamlyn, T. Mander, Y. Cungang, and R. Cheung. *Role-based model security access control for smart power-grids computer networks*. in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*. 2008.

[16] Crnkovic, I., S. Sentilles, A. Vulgarakis, and M.R. Chaudron, *A classification framework for software component models.* Software Engineering, IEEE Transactions on, 2011. **37**(5): p. 593-615.

[17] Cysneiros, L.M. and J.C. Sampaio do Prado Leite, *Non-functional requirements: from elicitation to modelling languages*, in *Proceedings of the 24th International Conference on Software Engineering, 2002. ICSE 2002*. 2002. p. 699-700.

[18] Dybå, T. and T. Dingsøyr. *Strength of evidence in systematic reviews in software engineering*. in *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*. 2008. ACM.

[19] Eby, M., J. Werner, G. Karsai, and A. Ledeczi. *Integrating security modeling into embedded system design*. in *Engineering of Computer-Based Systems, 2007. ECBS'07. 14th Annual IEEE International Conference and Workshops on the*. 2007. IEEE.

[20] Fernandez, E.B. and M.M. Larrondo-Petrie. *Designing secure SCADA systems using security patterns*. in *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*. 2010. IEEE.

[21] Fletcher, K.K. and X. Liu. *Security requirements analysis, specification, prioritization and policy development in cyber-physical systems*. in *Secure Software Integration & Reliability Improvement Companion (SSIRI-C), 2011 5th International Conference on*. 2011. IEEE.

[22] Forbes, L., H. Vu, B. Udrea, H. Hagar, X.D. Koutsoukos, and M. Yampolskiy. *SecureCPS: Defending a nanosatellite cyber-physical system*. in *SPIE Defense+ Security*. 2014. International Society for Optics and Photonics.

[23] Forrester, *Predictions 2016: Cybersecurity Swings To Prevention*. November 2015, Forrester.

[24] France, R., I. Ray, G. Georg, and S. Ghosh, *Aspect-Oriented Approach to Early Design Modelling*, in *IEE Proceedings - Software*. 2004, IEEE Computer Society. p. 173-185.

[25] Gunes, V., S. Peter, T. Givargis, and F. Vahid, *A Survey on Concepts, Applications, and Challenges in Cyber-Physical Systems.* KSII Transactions on Internet & Information Systems, 2014. **8**(12).

[26] Hahn, A. and M. Govindarasu, *Model-based intrustion detection for the smart grid (MINDS)*, in *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*. 2013, ACM: Oak Ridge, Tennessee, USA. p. 1-4.

[27] Hartmann, T., F. Fouquet, J. Klein, G. Nain, and Y. Le Traon, *Reactive security for smart grids using models@ run. time-based simulation and reasoning*, in *Smart Grid Security*. 2014, Springer. p. 139-153.

[28] Horvath, I. and B.H. Gerritsen. *Cyber-physical systems: Concepts, technologies and implementation principles*. in *Proceedings of TMCE*. 2012.

[29] Jauhar, S., C. Binbin, W.G. Temple, D. Xinshu, Z. Kalbarczyk, W.H. Sanders, and D.M. Nicol. *Model-Based Cybersecurity Assessment with NESCOR Smart Grid Failure Scenarios*. in *Dependable Computing (PRDC), 2015 IEEE 21st Pacific Rim International Symposium on*. 2015.

[30] Jensen, J.C., D.H. Chang, and E.A. Lee. *A model-based design methodology for cyber-physical systems*. in *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*. 2011. IEEE.

[31] Karnouskos, S. *Stuxnet worm impact on industrial cyber-physical system security*. in *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*. 2011. IEEE.

[32] Khaitan, S.K. and J.D. McCalley, *Design techniques and applications of cyberphysical systems: A survey.* Systems Journal, IEEE, 2015. **9**(2): p. 350-365.

[33] Kitchenham, B., *Guidelines for performing systematic literature reviews in software engineering.* EBSE Technical Report, 2007.

[34] Kitchenham, B.A., D. Budgen, and O.P. Brereton, *Using mapping studies as the basis for further research–a participant-observer case study.* Information and Software Technology, 2011. **53**(6): p. 638-651.

[35] Knirsch, F., D. Engel, C. Neureiter, M. Frincu, and V. Prasanna, *Privacy Assessment of Data Flow Graphs for an Advanced Recommender System in the Smart Grid*, in *Information Systems Security and Privacy*. 2015, Springer. p. 89-106.

[36] Lemaire, L., J. Lapon, B. De Decker, and V. Naessens. *A SysML extension for security analysis of industrial control systems*. in *Proceedings of the 2nd International Symposium on ICS & SCADA Cyber Security Research 2014*. 2014. BCS.

[37] Lucio, L., Q. Zhang, P.H. Nguyen, M. Amrani, J. Klein, H. Vangheluwe, and Y. Le Traon, *Advances in Model-Driven Security*. Advances in Computer, ed. A. Hurson and A. Memon. 2014: Elsevier.

[38] McDowell, M. *Understanding Denial-of-Service Attacks*. 2009 June 2016]; Available from: https://www.us-cert.gov/ncas/tips/ST04-015.

[39] McGraw, G., *Software security: building security in*. Vol. 1. 2006: Addison-Wesley Professional.

[40] Microsoft. *Security Development Lifecycle*. Available from: https://www.microsoft.com/en-us/sdl/.

[41] MIT. *Alloy Analyzer*. Available from: http://alloy.mit.edu.

[42] Mitchell, R. and I.-R. Chen, *A survey of intrusion detection techniques for cyber-physical systems.* ACM Computing Surveys (CSUR), 2014. **46**(4): p. 55.

[43] Mosterman, P.J. and J. Zander, *Cyber-physical systems challenges: a needs analysis for collaborating embedded software systems.* Software & Systems Modeling, 2016: p. 1-12.

[44] Nasr, P.M. and A.Y. Varjani. *Petri net model of insider attacks in SCADA system*. in *2014 11th International ISC Conference on Information Security and Cryptology, ISCISC 2014*. 2014.

[45] Neureiter, C., G. Eibl, D. Engel, S. Schlegel, and M. Uslar, *A concept for engineering smart grid security requirements based on SGAM models.* Computer Science-Research and Development, 2014: p. 1-7.

[46] Nguyen, P.H., J. Klein, Y. Le Traon, and M.E. Kramer, *A Systematic Review of Model-Driven Security*, in *Software Engineering Conference (APSEC, 2013) 20th Asia-Pacific*. 2013. p. 432-441.

[47] Nguyen, P.H., M. Kramer, J. Klein, and Y.L. Traon, *An extensive systematic review on the Model-Driven Development of secure systems.* Information and Software Technology, 2015. **68**: p. 62 - 81.

[48] Nielsen, C.B., P.G. Larsen, J. Fitzgerald, J. Woodcock, and J. Peleska, *Systems of Systems Engineering: Basic Concepts, Model-Based Techniques, and Research Directions.* ACM Computing Surveys (CSUR), 2015. **48**(2): p. 18.

[49] Oates, R., F. Thom, and G. Herries. *Security-aware, model-based systems engineering with SysML*. in *Proceedings of the 1st International Symposium on ICS & SCADA Cyber Security Research 2013*. 2013. BCS.

[50] Orojloo, H. and M.A. Azgomi. *A method for modeling and evaluation of the security of cyber-physical systems*. in *2014 11th International ISC Conference on Information Security and Cryptology, ISCISC 2014*. 2014.

[51] Pedroza, G., L. Apvrille, and D. Knorreck. *Avatar: A sysml environment for the formal verification of safety and security properties*. in *New Technologies of Distributed Systems (NOTERE), 2011 11th Annual International Conference on*. 2011. IEEE.

[52] Petersen, K., S. Vakkalanka, and L. Kuzniarz, *Guidelines for conducting systematic mapping studies in software engineering: An update.* Information and Software Technology, 2015. **64**: p. 1-18.

[53] Rajkumar, R.R., I. Lee, L. Sha, and J. Stankovic. *Cyber-physical systems: the next computing revolution*. in *Proceedings of the 47th Design Automation Conference*. 2010. ACM.

[54] Ruiz, J.F., R. Harjani, A. Mana, V. Desnitsky, I. Kotenko, and A. Chechulin. *A methodology for the analysis and modeling of security threats and attacks for systems of embedded components*. in *Parallel, Distributed and Network-Based Processing (PDP), 2012 20th Euromicro International Conference on*. 2012. IEEE.

[55] Saadatmand, M., A. Cicchetti, M. Sjödin, and T. Leveque, *Managing Timing Implications of Security Aspects in Model-Driven Development of Real-Time Embedded Systems.* International Journal On Advances in Security, 2012. **5**(3/4): p. 68-80.

[56] Sendall, S. and W. Kozaczynski, *Model transformation: the heart and soul of model-driven software development.* Software, IEEE, 2003. **20**(5): p. 42-45.

[57] Sridhar, S., A. Hahn, and M. Govindarasu, *Cyber–physical system security for the electric power grid.* Proceedings of the IEEE, 2012. **100**(1): p. 210-224.

[58] Suleiman, H., I. Alqassem, A. Diabat, E. Arnautovic, and D. Svetinovic, *Integrated smart grid systems security threat model.* Information Systems, 2015. **53**: p. 147-160.

[59] Suleiman, H. and D. Svetinovic, *Evaluating the effectiveness of the security quality requirements engineering (SQUARE) method: a case study using smart grid advanced metering infrastructure.* Requirements Engineering, 2013. **18**(3): p. 251-279.

[60] Tabrizi, F.M. and K. Pattabiraman. *A model for security analysis of smart meters*. in *Proceedings of the International Conference on Dependable Systems and Networks*. 2012.

[61] Tabrizi, F.M. and K. Pattabiraman. *A model-based intrusion detection system for smart meters*. in *Proceedings - 2014 IEEE 15th International Symposium on High-Assurance Systems Engineering, HASE 2014*. 2014.

[62] Ur-Rehman, O. and N. Zivic, *Secure Design Patterns for Security in Smart Metering Systems*, in *9th IEEE European Modelling Symposium on Mathematical Modelling and Computer Simulation*. 2015, IEEE: Madrid.

[63] Vasilevskaya, M., L.A. Gunawan, S. Nadjm-Tehrani, and P. Herrmann, *Integrating security mechanisms into embedded systems by domain-specific modelling.* Security and Communication Networks, 2014. **7**(12): p. 2815-2832.

[64] Vasilevskaya, M. and S. Nadjm-Tehrani, *Quantifying Risks to Data Assets Using Formal Metrics in Embedded System Design*, in *Computer Safety, Reliability, and Security*. 2015, Springer. p. 347-361.

[65] Vu, A.H., N.O. Tippenhauer, B. Chen, D.M. Nicol, and Z. Kalbarczyk, *CyberSAGE: a tool for automatic security assessment of cyber-physical systems*, in *Quantitative Evaluation of Systems*. 2014, Springer. p. 384-387.

[66] Wang, E.K., Y. Ye, X. Xu, S.-M. Yiu, L.C.K. Hui, and K.-P. Chow. *Security issues and challenges for cyber physical system*. in *Proceedings of the 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing*. 2010. IEEE Computer Society.

[67] Wang, W. and Z. Lu, *Cyber security in the Smart Grid: Survey and challenges.* Computer Networks, 2013. **57**(5): p. 1344-1371.

[68] Wasicek, A., P. Derler, and E. Lee. *Aspect-oriented modeling of attacks in automotive Cyber-Physical Systems*. in *Design Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE*. 2014. IEEE.

[69] Weigert, T. and F. Weil. *Practical experiences in using model-driven engineering to develop trustworthy computing systems*. in *Sensor Networks, Ubiquitous, and Trustworthy Computing, 2006. IEEE International Conference on*. 2006. IEEE.

[70] Wieringa, R., N. Maiden, N. Mead, and C. Rolland, *Requirements engineering paper classification and evaluation criteria: a proposal and a discussion.* Requirements Engineering, 2006. **11**(1): p. 102-107.

[71] Wohlin, C. *Guidelines for snowballing in systematic literature studies and a replication in software engineering*. in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. 2014. ACM.

[72] Wohlin, C. and R. Prikladnicki, *Systematic literature reviews in software engineering.* Information and Software Technology 2013. **55**(6): p. 919 - 920.

[73] Yampolskiy, M., P. Horvath, X.D. Koutsoukos, Y. Xue, and J. Sztipanovits. *Systematic analysis of cyber-attacks on CPS-evaluating applicability of DFD-based approach*. in *Resilient Control Systems (ISRCS), 2012 5th International Symposium on*. 2012. IEEE.

[74] Yampolskiy, M., P. Horváth, X.D. Koutsoukos, Y. Xue, and J. Sztipanovits, *A language for describing attacks on cyber-physical systems.* International Journal of Critical Infrastructure Protection, 2015. **8**: p. 40-52.

[75] Zafar, N., E. Arnautovic, A. Diabat, and D. Svetinovic, *System security requirements analysis: A smart grid case study*. Systems Engineering, 2014. **17**(1): p. 77-88.

[76] Zheng, X., C. Julien, M. Kim, and S. Khurshid, *On the state of the art in verification and validation in cyber physical systems*. The University of Texas at Austin, The Center for Advanced Research in Software Engineering, Tech. Rep. TR-ARiSE-2014-001, 2014.