Security in Model Driven Development: A Survey

Jostein Jensen
Department of Computer and Information Science
Norwegian University of Science and Technology
Trondheim, Norway
Email: jostein.jensen@idi.ntnu.no

Martin Gilje Jaatun
Department of Software Engineering, Safety and Security
SINTEF ICT
Trondheim, Norway
Email: Martin.G.Jaatun@sintef.no

Abstract—Model driven development (MDD) is considered a promising approach for software development. In this paper the results of a systematic survey is reported to identify state-of-the-art within the topic of security in model driven development, with a special focus on finding empirical studies. We provide an introduction to the major secure MDD initiatives, but our survey shows that there is a lack of empirical work on the topic. We conclude that better standardisation initiatives and more empirical research in the field is necessary before it can be considered mature.

I. INTRODUCTION

Model driven development (MDD) has been considered a promising approach to software development since its introduction about a decade ago. The Object Management Group (OMG)\(^1\) is the most prominent standardisation body within the MDD domain, and has developed a framework for model driven development called Model Driven Architecture (MDA). Improved portability, platform independence and cross-platform interoperability are among keywords used by OMG to describe the benefits of using a model driven software development process.

Figure 1 shows the MDA software development lifecycle as it is depicted by Kleppe et al. [1]. The ovals to the left represent generic software development phases, while the squares to the right represent artefacts produced in an MDA context. Artefacts developed during the requirements phase and used for analysis are often referred to as Computational Independent Models (CIM). Platform independent models (PIM) are abstract representations of the system to be built, and independent of any implementation technology. PIMs are transformed, preferably automatically using tool support, to Platform Specific Models. These are specific to the technology that will be used to realise future systems. Continuing the MDA lifecycle, PSMs are transformed into code. Since PSMs are close to the technology, this transformation is by some considered to be straightforward [1].

PIMs represent low-level system designs and as such constitute an important part of a system’s documentation (while still providing important abstractions). The layering between platform independent models, platform specific models and code are the key to solve problems related to portability, platform independence and interoperability.

\(^1\)http://www.omg.org/mda/executive_overview.htm

In traditional software development, security aspects are often considered late in the development lifecycle, if they are considered at all [2]. However, the cost of eliminating security flaws increases by magnitudes the later they are discovered and fixed [3]. A good recommendation has therefore been to include security aspects from the very start of software projects [4]. The Microsoft Security Development Lifecycle [5] and McGraw’s touchpoints [6] illustrate how security activities can be included in every phase of a software project.

With its focus on high-quality design in early development phases through PIM modelling, MDD/MDA should be a perfect development framework to include security aspects in design models from the very start of a project. Consistent and sound security solutions throughout the entire application could be the result.

The remainder of this paper is organised as follows: In Section II we present our research questions, followed by a description of our research method in Section III. We present our results in Section IV, and discuss our findings in Section V. Section VI concludes the paper.

II. RESEARCH QUESTIONS

This paper reports results related to a systematic survey that was carried out in order to learn how scientific communities deal with security in model driven development. The study aims to answer the following research questions:

- **RQ1:** What are the major scientific initiatives describing automatic code generation from design models within the context of security in MDD?
- **RQ2:** What empirical studies exist on the topic 'security within MDD/MDA?'
- **RQ3:** What are the strengths of the evidence showing that security aspects successfully can be modeled as an inherent property and transformed to more secure code?

III. METHOD

A systematic literature review approach is used as research method leading to the results presented in this paper [7]. This method requires rigor with respect to planning, conducting, and reporting the review. The aim of this systematic survey was to identify scientific literature that could provide answers to our research questions listed in the previous section.
A. Identification of research

The starting point for the survey is a research protocol where the research questions and the search strategy are defined. To support the paper selection process, the protocol also specifies inclusion and exclusion criteria.

A rigorous and comprehensive search is key to identify all the relevant scientific literature. Both sources for scientific literature and search phrases were specified prior to the search.

We used four online databases for scientific literature to search for studies:

- IEEE Xplore
- ACM Digital Library
- ISI Web of Knowledge
- Compendex

According to experiences made by Dybå et al. [8], this should be sufficient to find relevant literature within the information systems field. The use of other databases will lead to duplicate findings, and as such, lead to extra work. For each of these databases we used the following search phrases and keywords:

1) "model driven development"
2) "model driven architecture"
3) MDD
4) MDA
5) Security

These were combined as follows: (1 OR 2 OR 3 OR 4) AND 5.

The searches were performed March 12, 2010, meaning that scientific literature indexed up until then are included within this study. The search resulted in a total of 2844 titles that needed to be evaluated based on title, abstract and content.

B. Selection of primary studies

All references and abstracts were imported to the reference tool EndNote. The next step was to exclude papers based on titles. All titles that clearly did not treat the wanted topics were filtered out. After this process a total of 366 studies remained. The following task was to read through the abstracts of these papers and evaluate whether they were relevant or not. For both these steps the following exclusion criteria were used:

- Exclude everything that is clearly not related to model driven software development.
- Exclude everything that clearly not concerns both model driven development and security research.

122 papers remained after reading the abstracts, and these papers were all read to make a final evaluation whether they should be part of our primary studies or not. This evaluation resulted in 56 remaining papers. A last exclusion criterion was used for the purpose of this paper in order to answer RQ1:

- Exclude studies by authors and research groups who have published 3 or fewer papers on the topic.

122 papers remained after reading the abstracts, and these papers were all read to make a final evaluation whether they should be part of our primary studies or not. This evaluation resulted in 56 remaining papers. A last exclusion criterion was used for the purpose of this paper in order to answer RQ1:

- Exclude studies by authors and research groups who have published 3 or fewer papers on the topic.

This exclusion criterion can give a somewhat inaccurate view of the current state, as some important initiatives may be treated and enhanced by different research groups. For the purpose of this paper, it is however considered sufficient to give a rough idea about the current state. With this last exclusion criterion the number of papers to include as primary studies in this report was limited to 30.

C. Quality assessment, classification and synthesis

RQ2 and RQ3 can only be answered with a scientific validity if empirical studies following a rigorous research protocol on the topic are found. However, within the topic of model driven development and security, this study shows that none empirical studies seems to exist. Within this paper we therefore give a short introduction to the included papers considered for answering RQ1. Studies are grouped based on the originating research groups. A qualitative reflection about how MDD and security is covered in existing research works is given at the end of this paper.
IV. RESULTS

Our survey identified 5 research approaches which will be described in the following.

A. Model Driven Security

One of the earliest initiatives for including security in model driven architecture came from Basin et al. [9]. Their solution, called model driven security, is a specialisation of the MDA approach. Security models are integrated with what Basin et al. call UML process models, and the combined models are transformed into executable systems with integrated security infrastructures. The focus of their work is to include access control constraints based on role based access control (RBAC) in design models. A security metamodel for expressing RBAC properties in UML is given, and this UML extension is called SecureUML. In [10] Basin and Doser give a more detailed description of the Model Driven Security approach, while Clavel et al. [11] build on this work to gain practical experience with the approach.

B. SECTET

In [12] Alam et al. describe an approach to specify role-based access control policies for web services using the Object Constraint Language (OCL). OCL was initially a language extension of UML and is used to ensure a platform independent specification of access control policies. This work is used and extended by Breu et al. [13] who show how security can be built into web service-based systems supporting inter-organisational workflows. To model inter-organisational workflows they specify three model levels: global workflow model, local workflow model and interface model. The global workflow models show an abstract view of interactions between autonomous organisations, the local workflow models show intra-organisational workflows within each organisation, and the interface models present the services offered by each component in the system. OCL is used together with the interface models to describe access control constraints for operations/services provided by a web service. The same team builds on these concepts in [14] where the focus is to integrate security into the global workflow model. They use OCL-like expressions to assign security qualities such as confidentiality and integrity to data sent between actors.

The research team behind the above mentioned reports [12], [13], [14] has built on these results and come up with a model driven security framework called SECTET. Three software engineering paradigms are combined in this framework [15]: Model Driven Architecture as methodical concept, Service Oriented Architecture as architectural paradigm, and web services as technical standard. The three model levels described above is kept, and the OCL security policy definitions are refined into an OCL-based language they call SECTET-PL. Alam et al. [16], [17] present the SECTET framework with a focus on integrating access control policies in the interface models. They give a detailed description on how they specify dynamic access control constraints using SECTET-PL, and how these policy rules are combined with UML models at the interface level. In [18] SECTET-PL is used to describe how delegation rights in service-oriented architectures can be implemented, and in [19] and [20] SECTET is presented in a trust management perspective.

While the early reports [12], [13], [14] only were at the idea phase, [16] describes the whole tool chain to carry out model-to-model transformation and model-to-code transformation. They define UML meta-models for their concepts to formalise the modelling process sufficiently to allow tool-supported transformations, and Hafner et al. [21] focus on using the OMG transformation specification Meta Object Facility Query/View/Transformation (MOF-QVT) to formalise transformation rules.

Fernandez-Medina et al. [22] describe the SECTET framework to be one of the most complete frameworks to integrate security engineering with Model Driven Architecture.

C. Secure development of Data Warehouses

Data warehouses (DW) are repositories where enterprises electronically can store data from their various business systems. This is done to facilitate reporting and analysis of the data. Often data is “...extracted from multiple heterogeneous, autonomous, and distributed sources of information” [23]. Single data elements in the repository can be sensitive, but also the total amount of business information collected soon becomes business sensitive. Soler et al. [23] therefore argue that security engineering must be included from the earliest phases of development of such systems. Soler and his colleagues [23], [24], [25] argue that MDA is a well suited development framework to create DW solutions, but with the disadvantage that the MDA framework does not include mechanisms to sufficiently express security requirements, and as such perform a transformation from PIM to PSM. In their work based on the UML modelling language they show how they use UML profiles and model a security enriched PIM meta-model for the DW domain. In their framework, they also provide a set of QVT transformation rules so that PIMs can be transformed and mapped to concepts in a security enriched PSM meta-model that they also have defined. In addition to the security concepts defined in the two meta-models, dynamic security rules, such as audit and authorization rules can be added to the model using the OCL language. While Soler et al. focused on PIM to PSM transformations, Blanco et al. [26], [27] build on this work and demonstrate with a prototype that it is feasible to go all the way in the MDA lifecycle, from secure PIMs to secure PSM to code with security properties, in order to build secure data warehouses. Soler et al. supplement this work in [28].

The framework for development of secure data warehouses are further extended in [29] and [30]. In these works the authors build on the i* modelling language, which is designed

http://www.omg.org/spec/QVT/1.1/Beta2/

A definition of the Data Warehouse concept can be found at: http://searchsqlserver.techtarget.com/definition/data-warehouse

It may be argued that such mechanisms were never intended to be a part of MDA.
to support modelling of business requirements. *i*\(^a\) concepts are converted to a UML profile to fit the DW MDA approach, and some extensions are made to the original *i*\(^a\) concepts to be able to sufficiently express security requirements in the DW domain. This new *i*\(^a\) UML-profile supports elicitation of requirements at the business level, and is considered as being a CIM. Guidelines for transforming the business security requirements models to PIM are given to align the approach with MDA.

Blanco et al. [31] present an approach for modernising existing DWs by means of the above mentioned techniques for secure DW development. By going backwards in a reverse engineering style, they claim that code for existing DWs, presumably with insufficient security, can be analysed and converted to a PSM. This PSM is again transformed into a PIM, and finally a CIM. Now, the CIM can be analysed from a business perspective. Security requirements can be added, and then the new secure DW approach can be followed to get a more secure DW with the same functionality as it had before modernisation.

To bring the secure DW MDA approach closer to completion, Trujillo et al. [32] define an engineering process to support the framework shown in Figure 2. This paper defines the process that starts with *i*\(^a\)-based CIM models, which are transformed into secure PIMs, PSMs and code through transformation T1 to T3. It shows that security can be included from the very beginning of a project by using an MDA approach.

D. Security in business process models

Rodriguez et al. [33], [34] present initial ideas on how UML 2.0 activity diagrams, which are used to model business processes, can be enriched to include security properties. The authors claim that the advantage of including security in the business process modelling stage is that this important aspect then can be included from the very beginning of a software development project, and that a business analyst’s considerations about security can be captured. They define a UML profile consistent with OMG MOF, similar to the ideas of Hafner et al. [21]. A graphical notation to represent security requirements is added to the activity diagram notation.

In [35] the same authors suggest how the business process models, which they consider to be CIMs, can be transformed into use case models, which they consider to be PIMs. The transformation process is based on OMGs QVT specification, checklists and refinement rules. The feasibility of the approach is demonstrated through a prototype tool [36]. Use case models are often the starting point in software development projects where they are used to capture functional requirements. With this work, functional security requirements can be visually illustrated from the start in these models.

E. Secure smart card application development

Moebius et al. [37], [38], [39] use a model driven approach, which they call secureMDD, to develop security critical applications for smart cards. Their illustrating case is the development of an application that can be used for payment. From the PIMs they design, a transformation to three new model types is made: to card PSM, to terminal PSM and to a formal PSM. The two first model types define the functionality on, and interaction between the payment card and the terminal in which the card is used. The latter is a formal security specification of their models that can be analysed to determine the correctness with respect to security of their models.

Moebius and her colleagues emphasise the importance of both modelling static and dynamic aspects of the application. UML is the preferred modelling language in their approach.

The secureMDD approach is introduced in [37], and the approach to go from PIM to PSM to code is specified in more detail in [38]. Class diagrams are used to model an application’s static view, while sequence and activity diagrams are used for modelling of dynamic aspects. The transformation from PIMs to formal specifications is shown in [39].

V. Discussion

The existing papers on the topic can be categorized as lessons learned/experience reports where approaches at best are demonstrated by implementing prototypes. They provide little evidence to prove that the final code is more secure or better than what it would have been if another development approach had been used. The contribution that comes closest to being an empirical study is the paper written by Clavel et al. [11]. They provide an experience report where the MDS approach defined by Basin et al. (see section IV-A) has been tested in an industrial setting. Their feedback on the approach is quite optimistic, and with respect to MDS their major findings are:

- "The security design models integrate security models with system design models, remaining at the same time technology independent, reusable, and evolvable."
- "The security design models are understandable by those familiar with the UML-notation."
- The security-enhanced models were “expressive enough to model the access control policy defined in the original requirements document” provided by their clients.

This seems promising, but there are still several challenges that should be adressed in the coming years. Some of the promises of MDD/MDA are that the approach will ensure portability, platform independence and cross-platform interoperability. However, the studies included in this paper all explain different approaches for including security into the modelling languages and the processes they use. Since it is recognised that security modelling is not part of any standardisation initiatives for MDD, e.g. MDA, researchers define their own extensions to existing modelling languages to model the security aspects they need for their projects. An example of this is the use of OCL, which is the standardised UML constraint language used as starting point for specifying dynamic security aspects in the two most complete MDA frameworks: SECTET and secure DW. Both research teams found limitations with respect to modelling security constraints in the OCL language. Consequently, they started adapting it.
In the SECTET framework, the SECTET-PL was the resulting constraint language used, and in the DW design they extended a DW UML profile in order to better integrate concepts from the OCL expressions into their models. In general, standardisation initiatives exist with the purpose of encouraging the development of interoperable systems, so when standards are adapted and extended in different ways by different research teams it can be questioned whether final systems really will be interoperable and portable and so on.

McDermott [40] argues that one topic not sufficiently covered within security modelling, is related to modelling of security protocols. Moebius and her colleagues treat this in their approach for secure smart card application development. However, they do not follow a standardised MDD approach such as the MDA framework. At the same time it can be questioned whether the descriptions of their approach is sufficient to reconstruct their transformations from PIMs expressing protocol information to PSMs and then code. Thus, McDermott’s point still seems to be valid.

A key ingredient in MDD is the transformation rules guiding conversion from PIM to PSM to code. Based on the papers included in this study, the transformation rule development seems like a complex task, which requires a lot of expertise both with respect to the used development approach and technology platforms. This raises questions whether the team of security experts responsible for analysing security needs and requirements, also need to be experts on the modelling approach. If a transformation rule is flawed in a sense that it does not correctly transform a security requirement/model to code, then the whole system’s security can be compromised. Security experts should therefore also be able to evaluate the quality of transformation rules in all parts of the transformation chain to successfully benefit from the promises of security in MDD. Unfortunately, the situation seems to be that development teams and security teams often are separated, and that the real security experts usually do not themselves develop software [2]. This situation must be changed if high-quality secure code is going to be produced in an MDD context with automated code generation.

There is one important topic related to security that has not been discussed in the papers identified in this study; the possibility to model input validation constraints. Data sent to interfaces should be validated before they are accepted. Both the length and type of data must be checked in order to avoid security vulnerabilities related to injection attacks. To date, these types of vulnerabilities are the most prevalent security flaws in existing web applications. It should be possible to include modelling of input validation constraints in order to eliminate injection attack threats from the start of software development, similar to modelling of access control constraints.

A. Excluded Studies

There have been significant initiatives on topics related to this study that have been excluded due to RQ1 and the exclusion criteria used for the purpose of selecting primary studies. A notable example is UMLsec, an extension of UML supporting secure systems development [41]. Security requirements such as confidentiality, integrity and authenticity can be modelled in UML diagrams through the extension mechanisms stereotypes and tags. Modeling with UMLsec and analysis of industrial systems using this approach is even tested in industrial projects [42] [43] [44]. However, even though UMLsec is an important contribution to security engineering research in general and in the core of security in model driven development, papers on this topic were excluded due to our focus on automatic code generation.

Another topic not dealt with in this study is aspect oriented modeling. In aspect oriented modeling crosscutting concerns for an application, or aspects, are treated separately. Each aspect is then modelled and, by tool support, woven together into the final product. Examples of what an aspect might be include security, mobility and availability. Aspect oriented modelling papers were excluded since security was not treated specifically, but as one of several aspects. Still, we recognize that this approach may be worth looking into in future studies.

In the past there have been attempts to identify empirical research on the wide topic of model driven development. The systematic survey performed by Haug [45] returned a total of 21 papers, but this was only 2.2 % of the studies from the initial search; none with special focus on security. There were, however, limitations in this study with respect to sources used to find relevant literature; only selected journals and conference proceedings were searched. One of the key objectives of the review presented in this paper was to identify empirical

---

9 http://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project
studies on the topic of security in MDD, which is a narrow field compared to what Haug presented. A search strategy with a wider scope with respect to publication databases was used in hope of finding relevant literature despite the findings by Haug. However, the observations made in our study (including the studies that are not presented in this paper) indicate that such empirical studies do not exist for the topic of security in Model Driven Development.

B. Further Work

From the discussion above, the following paths for future research are identified:

- Empirical research should be performed to determine whether security successfully can be included properly in MDD/MDA to build more secure systems.
- Modelling of security should be included as a standardisation activity in the MDD frameworks, such as MDA.
- More research should be performed related to how security protocols can be modelled and transformed to final systems.
- Research should be performed to find an approach for modelling of input validation constraints.

Additionally, a follow-up of what is presented within this paper seems natural. Here we have presented an introduction to the major initiatives within the field of security in MDD. Future work must cover a deeper analysis, which includes evaluating the maturity of the presented approaches, to see if they are ready to be applied within an industrial setting. It is also worth studying the main differences and commonalities of each approach to determine to what extent their elements can be combined or reconciled. Finally, it is worth looking into a refinement of the research protocol, maybe widen the scope of the research questions and exclusion criteria, so that initiatives such as UMLsec and Aspect Oriented Modeling will be covered.

A more fundamental challenge, however, resides in the area of measuring code security, i.e., comparing two pieces of code to determine which is most “secure”. Current approaches are limited to counting the accumulated number of discovered bugs/flaws in a software product [46], or (reverse) modelling a given implementation and comparing it to an “ideal” model [42], [43], [44] – the latter approach assumes that the “ideal” model always will produce more secure code\(^{10}\), but unless you can measure the security property, there is no way to know for sure. It is not clear whether this problem is solvable, and we are not aware that anyone is currently working on it.

VI. Conclusion

In this paper we have presented state-of-the-art within security research in model driven development and identified the most comprehensive works. The study shows that there is a need for more empirical studies on the topic, and that standardisation is key to achieve the objectives of MDD/MDA, which are increased portability and interoperability.

REFERENCES


\(\text{10The authors of UMLsec state that “ Automated theorem provers and model checkers automatically establish whether the security requirements hold” [43], but this is no panacea if the security requirements themselves are flawed (or missing).}\)


