MODWIN Project:
Model Weaving for Insider Threat Prevention in Information Systems

LIG Laboratory

Involved teams: VASCO, SIGMA and CONVECS
Contacts:
Akram Idani (akram.idani@univ-grenoble-alpes.fr)
Mario Cortes Cornax (mario.cortes-cornax@univ-grenoble-alpes.fr)

1 Context

Information Systems (IS) may be vulnerable to several threats (such as malicious access and/or data theft) coming from inside the system, *i.e.* from trusted users who are already granted a legal access. This kind of threat is called "insider attack" in cyber-security, being well known to be difficult to tackle [8].

Studies done by IBM X-Force Research in the cyber-security landscape state that: "*In 2015, 60 percent of all attacks were carried out by insiders* [...] *and they resulted in substantial financial and reputational losses*". Indeed, the problem is beyond the access control frontier since it includes unpredictable human behaviours. To deal with these threats, existing industrial, academic and government studies elaborate human profiles and advocate for the use of surveillance systems [1, 2, 6]. Without being exhaustive, some of these profiles are:

- Curious persons who, without a malicious intention but without self-control too, get access to sensitive data or do some actions that are in contradiction with the company rules.
- Super-heroes who, in order to fix a problem or help someone, bypass the company policies believing that it may be useful or simply be approved.

Other profiles are established in the literature, like machiavellian, greedy, disgruntled, opportunistic, etc. Unfortunately, the eventuality of a breach of trust is difficult to predict in advance based on human-centric factors. On the one hand there is no certainty about a possible acting out, and on the other hand people surveillance must comply with privacy legislation, which makes it almost ineffective. Nonetheless, Information Systems (IS) together with their business logic and processes, provide useful knowledge allowing one to deal with the insider threat problem. In fact, based on the aforementioned studies it can be observed that insiders often do not have high computer skills (contrary to intruders), but they have a fine-grained knowledge about the IS procedures. The latters are mostly well-established and already protected via access control mechanisms. Hence, by being able to answer the question "*who has access to sensitive data and what kind of access is given?*", one cannot claim that the system is secure enough. A more secure-oriented question should be: "*is the user able to run a sequence of actions that may bring him from a prohibition to an authorization?*". The first question refers to static concerns, and it is widely addressed in Model-Driven Security (MDS) thanks to several access control models (SecureUML, UMLSec, etc). However, the second question remains open in MDS because it refers to behavioural features and the reachability of unwanted situations granting to the user misappropriated privileges.

2 Challenges and research directions

A direct consequence of the separation of concerns principle in IS modeling is that database definition, business processes and security models are often validated separately. Existing works in MDS are therefore stateless and they mostly validate these concerns without taking into account their intertwining. Furthermore, these concerns are often defined by different stakeholders, leading to several inconsistencies between the resulting artifacts. These issues open security breaches

that may favour insider threats. A well known insider attack is the one that occurred in 'Société Générale'. This attack resulted in a net loss of $7.2 billion to the bank[1]. The insider circumvented internal security mechanisms to place more than $70 billion in secret, unauthorized derivatives trades. Through authorized actions and acceptable processes, he was able to cover up operations he made on the market by introducing into the system fictive offsetting inverse operations, so that the unauthorized trades were not detected. To deal with this kind of threats in IS, we focus on two challenges:

- Capture relationships between the various IS concerns (data, security policies and processes). To this purpose we suggest to use dedicated domain-specific languages and formal methods in order to define their respective semantic domains. On this aspect, a preliminary work has been done in [4], where we proposed a first version of a weaving model as well as a prototype in order to map three languages supporting the aforementioned concerns.
- Propose automated dynamic analysis of reachability properties to establish that a system evolves as expected and that unwanted situations are not possible, which would prevent insider threats. On this propose, we suggest the integration of already developed tools in the LIG lab; especially, Meeduse[2] (by VASCO) and CADP[3] (by CONVECS).

3 Solutions to explore

The gap between data, security policies and business processes is generally due to the lack of alignment between their underlying models. In fact, the data model and the business process model are often designed and developed by different teams even if a Business Process Management System (BPMS) is used: the data management team, which is mainly interested in the database implementation, and the process management team, that is composed of analysts without particular software skills. The fact that the standard language for process modeling BPMN (Business Process Model and Notation) does not integrate the data model, accents this gap. Since these teams use their own models and methodologies, the resulting artifacts are unrelated and the interaction between data and processes remains informal. Thereby, this issue gave rise to a plethora of data-centric process modeling approaches, but they do not support security concerns. Furthermore, most of access control models and engines are fine grained and start by identifying sensitive data of an IS and the corresponding protection mechanisms, but they do not support business process modeling. These issues recall a more general challenge, which concerns model weaving in the MDE community. The idea of model weaving is to capture and reify relationships between heterogeneous model elements. This would lead to a so-called weaving model, whose elements represent relationships (*e.g.* equality, concatenation, equivalence, etc) and the related elements.

In our context, models refer to: data model (*e.g.* UML class diagram, E/R schema, domain ontology), security model (*e.g.* RBAC, ABAC) and business process model (*e.g.* BPMN, UML activity diagram). To support the integration of the underlying modeling languages and notations and favour a formal reasoning about their correctness, the current project builds on the Meeduse [3] language workbench. Meeduse is the only existing language workbench today that allows one to formally instrument a domain-specific language in order to prove its correctness. The tool has had several realistic applications and won two awards at the Transformation Tool Contest (TTC'19): best verification and audience award. Meeduse applies B specifications to define the semantics of a DSL and embeds ProB [7] (a model-checker of the B method) for the execution capabilities. The formalisation of the three concerns in Meeduse belongs to an ongoing collaboration between VASCO and SIGMA [4]. The two teams shared their skills in business process modeling (SIGMA) and model-driven security (VASCO) and co-supervised two M2 students about this topic, resulting in an international workshop publication [4]. The aim of the current project is, on the one hand, to strengthen this collaboration and advance the current results, and on the other hand, to provide practical and tool-assisted solutions for the extraction of security flaws in IS.

The weaving model leading to the alignment of the three concerns is under construction and is not yet formally defined. In fact, it is important to identify the required design strategies in order to guide the definition of access control rules within the various models. These strategies can be:

- extend a business process model by some security mechanisms without defining explicit data access control; or

[1] The New York Times. French Bank Says Rogue Trader Lost $7 Billion. January 2008.
[2] http://vasco.imag.fr/tools/meeduse/
[3] https://cadp.inria.fr

- define explicit links between the various models and then verify their conformity.

Several research works addressed the first strategy and proposed to extend the business processes modeling languages with security aspects. We believe that extending standard notations is difficult, from an adoption point of view, as it implies the modification of languages and tools that practitioners already use. It also requires new design methodologies that may change the current IS practices.

In this project, we advocate for the second strategy and propose to use business processes as a way to coordinate the execution of the other models. Indeed, we can assume that every model provides a set of operations (such as read/write operations, user connections, verification of permissions, etc) and that based on the model-weaving concept, a process model may act like an orchestrator of these operations. The formal definition of data models and security rules in B is straightforward and can be built on the previous works of the VASCO team [5]. However, business process models provide a different point of view that is process-centric and hence process algebra techniques are much more suitable. This approach has been investigated by the CONVECS team leading to the VBPMN tool. The latter translates business process models into LNT, which allows one to apply the CADP toolbox in order to verify various kinds of reachability properties.

The challenge is therefore to develop a technique, built on B and LNT for the definition and the verification of secure IS: B models would be extracted from data models and security policies, and LNT specifications would be extracted from business process models. Having these formal specifications, we propose to orchestrate B operations using LNT processes. Another alternative solution would be to translate the whole weaving model into LNT and favour the verification and animation of the resulting formal specification in Meeduse. The proposed approach covers the various concerns of an IS, which provides a formally defined framework for the identification of insider threats. Regarding tools, connections between CADP and Meeduse will be investigated, and generalized to other kinds of model weaving.

References

[1] Frank L. Greitzer. Insider Threats: It's the HUMAN, Stupid! In *Proceedings of the Northwest Cybersecurity Symposium*, NCS '19, New York, NY, USA, 2019. Association for Computing Machinery.

[2] Ivan Homoliak, Flavio Toffalini, Juan Guarnizo, Yuval Elovici, and Martín Ochoa. Insight Into Insiders and IT: A Survey of Insider Threat Taxonomies, Analysis, Modeling, and Countermeasures. *ACM Computing Surveys*, 52(2), 2019.

[3] Akram Idani. The B method meets MDE: review, progress and future. In Renata S. S. Guizzardi, Jolita Ralyté, and Xavier Franch, editors, *16th International Conference on Research Challenges in Information Science*, volume 446 of *LNCS*, pages 495–512. Springer, 2022.

[4] Akram Idani and Mario Cortes Cornax. Towards a model driven formal approach for merging data, access control and business processes. In Esther Guerra and Ludovico Iovino, editors, *ACM/IEEE 23rd International Conference on Model Driven Engineering Languages and Systems, Companion Proceedings*, pages 57:1–57:5. ACM, 2020.

[5] Akram Idani and Yves Ledru. B for modeling secure information systems - the b4msecure platform. In Michael J. Butler, Sylvain Conchon, and Fatiha Zaïdi, editors, *17th International Conference on Formal Engineering Methods*, volume 9407 of *LNCS*, pages 312–318. Springer, 2015.

[6] Markus Kont, Mauno Pihelgas, Jesse Wojtkowiak, Lorena Trinberg, and Anna-Maria Osula. *Insider Threat Detection Study.* The NATO Cooperative Cyber Defence Centre of Excellence, 2018.

[7] Michael Leuschel and Michael J. Butler. Prob: an automated analysis toolset for the B method. *Int. J. Softw. Tools Technol. Transf.*, 10(2):185–203, 2008.

[8] Christian W. Probst, Jeffrey Hunker, Dieter Gollmann, and Matt Bishop, editors. *Insider Threats in Cyber Security*, volume 49 of *Advances in Information Security.* Springer, 2010.