Information and Software Technology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Modeling Security and Privacy Requirements: a Use Case-Driven Approach

Phu X. Mai[a], Arda Goknil[*,a], Lwin Khin Shar[c], Fabrizio Pastore[a], Lionel C. Briand[a], Shaban Shaame[b]

[a] SnT Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg
[b] EverdreamSoft (EDS), Switzerland
[c] School of Computer Science and Engineering, Nanyang Technological University, Singapore

A B S T R A C T

Context: Modern internet-based services, ranging from food-delivery to home-caring, leverage the availability of multiple programmable devices to provide handy services tailored to end-user needs. These services are delivered through an ecosystem of device-specific software components and interfaces (e.g., mobile and wearable device applications). Since they often handle private information (e.g., location and health status), their security and privacy requirements are of crucial importance. Defining and analyzing those requirements is a significant challenge due to the multiple types of software components and devices integrated into software ecosystems. Each software component presents peculiarities that often depend on the context and the devices the component interact with, and that must be considered when dealing with security and privacy requirements. *Objective:* In this paper, we propose, apply, and assess a modeling method that supports the specification of security and privacy requirements in a structured and analyzable form. Our motivation is that, in many contexts, use cases are common practice for the elicitation of functional requirements and should also be adapted for describing security requirements. *Method:* We integrate an existing approach for modeling security and privacy requirements in terms of security threats, their mitigations, and their relations to use cases in a misuse case diagram. We introduce new security-related templates, i.e., a mitigation template and a misuse case template for specifying mitigation schemes and misuse case specifications in a structured and analyzable manner. Natural language processing can then be used to automatically report inconsistencies among artifacts and between the templates and specifications. *Results:* We successfully applied our approach to an industrial healthcare project and report lessons learned and results from structured interviews with engineers. *Conclusion:* Since our approach supports the precise specification and analysis of security threats, threat scenarios and their mitigations, it also supports decision making and the analysis of compliance to standards.

1. Introduction

Modern internet-based services like home-banking [1], music-streaming [2], food-delivery [3], and personal-training [4] are delivered through multi-device software ecosystems, i.e., software systems with components and interfaces that are executed on different types of devices including Web browsers, desktop applications, mobile applications, smart-TVs, and wearable devices. Most of the multi-device software ecosystems process private end-user data collected and stored by different devices, such as credit balance reported by banking applications, locations visited by end-users, and health status tracked by personal training applications. The adoption of multi-device software ecosystems augments security and privacy risks because of the presence of multiple attack surfaces (points at which security attacks can be executed), including malware that steals consumer and corporate data from smartphones [5] and Web applications that unintentionally expose confidential data [6]. Therefore, security and privacy have become a crucial concern in the development of software ecosystems, starting from requirements analysis to testing.

To identify the security requirements of a multi-device software ecosystem, it is necessary to take into consideration the characteristics of the specific service being developed and of the device types on which the service is going to be deployed. An example requirement of a home-banking smartphone service is that the user should automatically log off when the phone screen is locked to prevent phone thieves from accessing the bank account. This requirement is inappropriate for other types of services, e.g., personal training services which are used by runners and thus should be accessible without logging in, even after a screen

* Corresponding author.
E-mail addresses: xuanphu.mai@uni.lu (P.X. Mai), arda.goknil@uni.lu, goknil@svv.lu (A. Goknil), lkshar@ntu.edu.sg (L.K. Shar), fabrizio.pastore@uni.lu (F. Pastore), lionel.briand@uni.lu (L.C. Briand), shaban@everdreamsoft.com (S. Shaame).

https://doi.org/10.1016/j.infsof.2018.04.007
Received 9 November 2017; Received in revised form 19 February 2018; Accepted 16 April 2018

lock (which normally happens while running). Examples of device specific characteristics that impact on security requirements include Web applications running on dedicated servers that are always online and thus prone to brute force attacks via the network. Mobile applications, instead, are often idle or offline, but they usually run on a device that is potentially shared with malicious applications inadvertently installed by end-users. Such applications can steal private data if it is not properly protected (e.g., through encryption). Therefore, it is crucial to precisely model and analyze security and privacy requirements of such multi-device software ecosystems early in their development.

In this paper, we propose, apply, and assess a use case-driven modeling method that supports the specification of security and privacy requirements of multi-device software ecosystems in a structured and analyzable form. Use cases are one of the most common means adopted by software engineers to elicit requirements because they ease the communication between stakeholders [7]. Therefore, to achieve widespread applicability, the need for integrating security requirements with use case modeling warrants the development of a use case-driven, security requirements modeling method that is, in our context, tailored to the development of multi-device software ecosystems.

Considerable research has been devoted to eliciting and analyzing security requirements using various forms of use cases (e.g., abuse cases [8,9], security use cases [10], and misuse cases [11–14]). However, the applicability of these approaches in the context of security and privacy requirements modeling for multi-device software ecosystems shows limitations with respect to (1) their support for explicitly specifying various types of security threats (a security threat is a possible event that exploits a vulnerability of the system to cause harm), (2) the definition of threat scenarios (a threat scenario is a flow of events containing interactions between a malicious actor and system to cause harm), and (3) the specification of mitigations for these threats.

These three features are essential in the type of business context we target where it is required to explicitly identify the threat scenarios that may affect important business operations in order to identify appropriate mitigation schemes and trade-offs between functional requirements and security and privacy concerns. It is also expected that such security requirements, specified in a structured and analyzable form, provide support for security testing, for example by helping with the identification of attack surfaces. In addition to specifying security threats, a common practice in many environments requires mitigation schemes to be documented for the stakeholders to demonstrate compliance with applicable security and privacy standards and regulations. However, existing approaches lack reusable templates to specify such mitigation schemes.

The goal of this paper is to address the above challenges by proposing a use case-driven, security requirements modeling method called *Restricted Misuse Case Modeling (RMCM)*, which adapts existing methods and extends them. In RMCM, we employ misuse case diagrams proposed by Sindre and Opdahl [13] to model security and privacy requirements in terms of use cases. Misuse cases describe attacks that may compromise use cases; security use cases specify how to mitigate such attacks. For eliciting security threats and threat scenarios in a structured and analyzable form, we adopt the Restricted Use Case Modeling method (RUCM) proposed in [15] to write use case specifications. RUCM is based on a template and restriction rules, reducing ambiguities and incompleteness in use cases. It was previously evaluated through controlled experiments and has shown to be usable and beneficial with respect to making use cases less ambiguous and more amenable to precise analysis and design [16–23]. However, since RUCM was not originally designed for modeling security and privacy requirements, we extend the RUCM template with new restriction rules and constructs, targeting the precise modeling of security threats. Further, we provide a template for mitigation schemes and three mitigation schemes that are pre-specified with standard and secure coding methods for mitigating common security threats. They can be readily used and revised as necessary.

In this paper, we focus on one important aspect of privacy: the security of personal data. More specifically, we support the modeling of requirements regarding three data protection goals [24]: confidentiality, integrity and availability. The definition of methods to model other data protection requirements (e.g., unlinkability, transparency, and intervenability) and to target privacy-related activities other than information processing (e.g., dissemination or collection [25]) is out of the scope of this paper (related work includes reports and regulations on data minimization, collection limitation, and purpose specification [26–29]).

Leveraging on the analyzable form of our models, RMCM employs Natural Language Processing (NLP) to report inconsistencies between a misuse case diagram and its RMCM specifications, and to analyze the compliance of such specifications against the provided RMCM templates. NLP is also used to identify and highlight the control flow leading to different threat scenarios and the steps in RMCM specifications that refer to interactions between malicious actors and the system. The latter provides security testers with information about attack surfaces on which security testing should focus. To summarize, the contributions of this paper are:

- RMCM, a security requirements modeling method supporting the precise and analyzable specification of security threats, threat scenarios, and their mitigations, in the context of use case driven development of multi-device software ecosystems;
- a practical toolchain, available at our tool website [30], including (1) a component that extends Papyrus [31] to support misuse case diagrams, (2) a component that extends IBM Doors [32] to support misuse case specifications and mitigation schemes in the RMCM templates, and (3) a component relying on NLP to detect inconsistencies among these artifacts;
- a case study demonstrating the applicability of RMCM in a realistic development context involving multiple service and software providers in the healthcare domain.

This paper is structured as follows. Section 2 introduces the context of our case study to provide the motivations behind RMCM. Section 3 describes the limitations of state-of-the-art approaches that we identified by concretely applying these approaches on our industrial case study. Section 4 discusses the related work. Section 5 provides an overview of RMCM. Section 6 focuses on the use case extensions in RMCM. In Section 7, we present our tool support. Section 8 reports on our industrial case study, from which we draw conclusions on the benefits and applicability of the proposed approach. Section 9 concludes the paper.

2. Context and motivation

The work presented in this paper is part of a European Union (EU) project, i.e., EDLAH2 [33], in the healthcare domain. The project brings academic institutions and software development companies together in a consortium to enhance the lifestyle of elderly people through a *gamification-based* approach. Gamification transforms activities that we are normally reluctant to do, e.g., exercising regularly, into a competition [34]. The objective of the EDLAH2 project is to provide a set of gamification-based services on mobile devices that engage and challenge clients (elderly people) to improve their physical, mental, and social activities.

To achieve this objective, the EDLAH2 consortium is developing a multi-device software ecosystem, i.e., a set of software components that can run on multiple types of systems and devices, which include mobile and wearable device applications (services). In EDLAH2, the mobile applications are used to incentivize elderly people to perform intellectual activities (e.g., solving logic-based games including Sudoku), while the wearable device applications are used to track physical