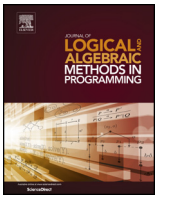
Journal of Logical and Algebraic Methods in Programming ••• (••••) •••–•••

Contents lists available at ScienceDirect

Journal of Logical and Algebraic Methods in Programming

www.elsevier.com/locate/jlamp



Formal methods for web security ☆

Michele Bugliesi, Stefano Calzavara *, Riccardo Focardi

Università Ca' Foscari Venezia, Italy

ARTICLE INFO

Article history:
Received 23 December 2015
Received in revised form 29 August 2016
Accepted 30 August 2016
Available online xxxx

Keywords:
Formal methods
Web security
Survey

ABSTRACT

In the last few years, many security researchers proposed to endow the web platform with more rigorous foundations, thus allowing for a precise reasoning on web security issues. Given the complexity of the Web, however, research efforts in the area are scattered around many different topics and problems, and it is not easy to understand the import of formal methods on web security so far. In this survey we collect, classify and review existing proposals in the area of formal methods for web security, spanning many different topics: JavaScript security, browser security, web application security, and web protocol analysis. Based on the existing literature, we discuss recommendations for researchers working in the area to ensure their proposals have the right ingredients to be amenable for a large scale adoption.

1. Introduction

The Web is now part of everyone's life and it constitutes the primary means of access to many useful services with strict security requirements. As a result, vulnerabilities on the web platform may enable vicious attacks with catastrophic consequences, ranging from economic losses, e.g., in the case of attacks against payment providers like PayPal, to privacy violations, e.g., in the case of improper disclosure of electronic health records. Security-critical services are more and more supplied online today and this increases the need of effective defenses for the web platform.

Unfortunately, it is well-known that protecting online services is not easy at all, given the intrinsic complexity of the Web. The web ecosystem is variegate and includes a large number of different components and technologies, hence the attack surface against web applications is incredibly large: security flaws in the web browser may expose authentication credentials and sensitive data stored in web pages; vulnerabilities of web protocols may break the confidentiality and the integrity of the communication session; and errors in the web application code may lead to the inclusion of malicious contents in otherwise trusted web pages. Even experienced web developers and security practitioners have a hard time at taming this complexity, leading to the proliferation of security breaches.

As it normally happens in computer science, when some kind of process is too error-prone, *formal methods* come to the rescue. In the last few years, many security researchers proposed to endow the web platform with more rigorous, analytical foundations. Their goal is designing models which allow for a precise reasoning on web security issues and developing effective tools to make the Web a safer place, relieving at least part of this burden from the shoulders of web developers and browser vendors. Given the complexity of the Web, however, research efforts in the area are quite scattered around many different topics and problems, and it is not easy to understand the import of formal methods on web security so far.

☆ Work partially supported by the MIUR projects CINA, ADAPT and Security Horizons.
* Corresponding author.
E-mail addresses: bugliesi@unive.it (M. Bugliesi), calzavara@dais.unive.it (S. Calzavara), focardi@unive.it (R. Focardi).

http://dx.doi.org/10.1016/j.jlamp.2016.08.006
2352-2208/

Please cite this article in press as: M. Bugliesi et al., Formal methods for web security, J. Log. Algebraic Methods Program. (2016), http://dx.doi.org/10.1016/j.jlamp.2016.08.006

One natural question is whether formal methods have been successful in this field or whether they can only be considered a theoretical exercise as of now: practical applications are important to showcase the effectiveness of formal methods at dealing with the problems mentioned above and encourage the web security community to synergise efforts with the formal methods community.

Through this survey, we make the following contributions:

1. we identify the most important, though occasionally underestimated, challenges which must be faced by researchers interested in investigating the application of formal methods to web security (Section 3);
2. we collect, classify and review existing proposals in the area of formal methods for web security, spanning many different topics: JavaScript security, browser security, web application security, and web protocol analysis. We underline the practical applications of the different solutions and we identify several success stories among them (Sections 4-7);
3. we discuss recommendations for researchers working in the area of formal methods for web security to ensure their proposals have the right ingredients to be amenable for a large scale adoption (Section 8).

1.1. Scope of the survey

In this survey, we review:

- models of common web technologies, like web browsers, and foundational studies on the semantics of scripting languages used by web developers;
- semantics-based tools for the verification and the enforcement of security properties on the web platform;
- alternative, provably sound designs of solutions aimed at replacing existing web technologies to improve their security.

Instead, given our declared goals, we do not review:

- novel security models and abstract proposals for the Web which have not been backed-up by an implementation and an (at least preliminary) on-field evaluation. The Web is a very heterogeneous and complex environment, hence it is impossible to assess the adequacy of new security mechanisms without any practical evaluation;
- tools and solutions which have not been formalised or proved correct with respect to a precise security definition, even if these proposals are loosely inspired by sound principles predicated in the formal methods literature, e.g., on type-safe programming languages or information flow control.

We also exclude from the present survey the rich research line on the verification of the TLS/SSL protocol used for secure communication on the Web. Though formal methods boast many success stories in this area, the topic would better fit a survey on protocol verification.

1.2. Organisation

At a high level, we observe that the proposals we survey can be divided in two main research lines:

(RL1) *security by construction*: some works advocate the usage of better languages and abstractions to make the Web a safer place. They typically recognise severe intrinsic limitations in the design of the current Web and propose a paradigm shift to improve it. These proposals are effective at solving the root cause of a security problem, but they typically require profound changes to existing web technologies and applications;

(RL2) *modelling, verification and enforcement*: some works propose models and algorithms to formalise and reason about the security of current web technologies. They devise solutions to make the Web a more secure place by exploiting the existing frameworks and standards at their best. These proposals may be sometimes sub-optimal in terms of effectiveness, but they do not impact too much on current web technologies.

These two research lines are thus largely complementary and equally important. The presentation in the next sections is based on this classification. When a formal model found successful practical applications we discuss them in a *Security Applications* paragraph.

1.3. Structure of the survey

Section 2 provides some background information about the web platform and web security in general. Section 3 discusses the main challenges in the application of formal methods to web security. Sections 4 and 5 overview formal methods for web security from the browser perspective: specifically, Section 4 focuses on JavaScript security, while Section 5 discusses other relevant work on browser security. Section 6 presents formal methods for securing web applications. Section 7 discusses formal models for web protocols, aimed at analysing both browsers and web applications, as well as their re-

Please cite this article in press as: M. Bugliesi et al., Formal methods for web security, J. Log. Algebraic Methods Program. (2016), http://dx.doi.org/10.1016/j.jlamp.2016.08.006