Available online at www.sciencedirect.com

ScienceDirect

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

Computers & Security

# Ontology for attack detection: An intelligent approach to web application security

Abdul Razzaq [a,*], Zahid Anwar [a], H. Farooq Ahmad [a,b], Khalid Latif [a], Faisal Munir [a]

[a] School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology, Islamabad, Pakistan

[b] College of Computer Sciences and Information Technology (CCSIT), King Faisal University, Alahssa 31982, Kingdom of Saudi Arabia

ARTICLE INFO

*Article history:*
Received 14 November 2013
Received in revised form
29 April 2014
Accepted 23 May 2014
Available online 12 June 2014

*Keywords:*
Web application security
Ontology based intelligent system
Semantic security
Cyber security
Information security

ABSTRACT

Conventional detection techniques struggle to keep up with the inherent complexity of web application design and hence the ever growing variety of attacks that can exploit it. Security frameworks modeled using an ontological approach are a promising new line of defense that can be highly effective in detecting zero day and sophisticated web application attacks because they can capture the context of the contents of information such as HTML pages or in-line scripts and have the ability to filter these contents by taking into consideration their consequences to the target applications. The goal of this article is to demonstrate how an ontology-engineering methodology may be systematically applied for designing and evaluating such security systems. A detailed ontological model is shown that caters to the generalized working of web applications, the underlying communication protocols and attacks. More specifically the proposed ontological model because it captures the context can not only detect HTTP protocol specification attacks but also helps focus only on specific portions of the request and response where a malicious script is possible. The model also captures the context of important attacks, the various technologies used by the hackers, source, target and vulnerabilities exploited by the attack, impact on system components and controls for mitigation. A comprehensive and best metrics suite for ontology evaluation has been used for assessing the quality of proposed model which includes correctness, accuracy, consistency, soundness, task orientation, completeness, conciseness, expandability, reusability, clarity, integrity, efficiency and expressiveness. The proposed model ranked well against the above mentioned metrics. Moreover a prototype attack detection system based upon the proposed model showed improved performance and detection rate and low rate of false positives while detecting OWASP's top ten listed web attacks.



* *Corresponding author.*
E-mail addresses: abdul.razzaq@seecs.nust.edu.pk (A. Razzaq), zahid.anwar@seecs.nust.edu.pk (Z. Anwar), farooq.ahmad@seecs.nust.edu.pk (H.F. Ahmad), khalid.latif@seecs.nust.edu.pk (K. Latif), faisal.munir@seecs.nust.edu.pk (F. Munir).
http://dx.doi.org/10.1016/j.cose.2014.05.005
0167-4048/

## 1. Introduction

The ever increasing growth in the number of cyber threats due to the popularity of web applications, is creating a strong security concern for the e-community. The latest Data Breach Investigative Report by Verizon (Wysopal, 2012) highlighted that 81% of hacking attacks were directed toward web applications, posing security risks to many organizations.

Various generic security controls such as signature-based firewalls, intrusion detection and prevention systems and encryption devices have been deployed, however their effectiveness against web-based threats is restricted, because of their extreme rigidness. For efficient mitigation of web based attacks the system should understand the context of the information contents to be processed and have the ability to filter the contents on the basis of their effect on the target application. Since it is difficult to capture the actual attack context in a traditional security solution, it has been previously proposed (Razzaq et al., 2014) that a security framework modeled using an ontological approach can be more effective in detecting zero day and sophisticated web application attacks. The current work describes the key conceptual differences between ontological and non-ontological security solutions with specific details on how an ontology-engineering methodology may be applied for designing and evaluating such systems.

Non ontological approach systems face a number of challenges, a few of which are outlined below:

- Most existing techniques are signature based, which maintain the syntactic representation of the attack. It is easy for an attacker to launch an attack by slight modification of this syntactical representation of the signature. One major challenge is the design of a system that represents the attack with a balanced abstraction to cater for similar variations of any particular attack.

- Current web application attack detection techniques are reactive; attacks are detected by frequently scanning the data and system logs; only being prevented if the exact signature of the specific attack is present and recognized by the system (otherwise the attack may not be detected and may compromise the security of the system). Thus it is necessary to design techniques that are proactive and provide the necessary measures to prevent exploitation of vulnerabilities that may damage the application.

- Behavior based IDSs (Intrusion Detection Systems) running at the application layer, that are not signature based, may detect new and previously unknown attacks. However, in these systems, a small deviation from the training data creates high false positives and false negatives thus making it a challenge to design a system that minimizes these to effectively detect zero day attacks without requiring the training data.

- Statistical techniques used in IDSs basically provide a workable solution for the network layer. This solution is not effective at the application layer because it focuses on the character distribution of the input and does not take into account its contextual nature.

- Some IDS solutions exist that provide signatures as blueprint for SQL Injection based attacks using malformed queries, depending on the application source code. Again similar to signature based detection these require modification when new source code is added for additional functionality.

- Capturing the context of input and output is again a challenging task and apart from application context, capturing the protocol context is also hard to achieve. For this propose ontological model of protocol, user request and corresponding response are formally modeled to address the issue. Automatic rule generation is also a fairly difficult and challenging task. Formally designing of semantic rules in the proposed models and inference capability for derivation of further rules, justifiably overcomes this difficulty.

### 1.1. *Ontology-based vs non-ontology based approach*

Ontology refers to the explicit specification of the conceptualization of a domain which captures its context (Interpretation of words in specific domain/situation). Mostly non-ontological approaches are based upon attack signatures. Normally the attack signatures are not generic in nature, using specific languages related to particular domains and depend upon specific environments and systems. Consequently, they lack extensibility and are also not suitable for communication in heterogeneous systems. Attack signatures often carry vague semantic knowledge and lack solid ground in any formal logic. Small variations in the business logic invalidate the signatures. Constantly updating of the signature is also a very tedious job.

Our ontological model caters to the generalized working of web applications, HTTP protocol, and attacks. This helps in improved detection for application level attacks if the security system works under the context of all of these. In short an ontology approach provides us with the following features as compared to other competing approaches.

Ontological model is flexible in defining any concept to the desired level of detail, is easily extendible and provides reasoning ability to reason over the instances of the data within the domain. Moreover ontological models can be shared and reused among the entities within the domain.

### 1.2. *Contributions*

The proposed ontology of attack and ontology of communication protocol provides a powerful construct to improve the detection capability of application level attacks. The proposed mechanism is a novel approach for employing the use of semantics in application layer security contrary to tradition signature based approaches. It has the following key contributions:

- Ontological model of communication protocol: The ontological model of HTTP captures the context of