Safeguarding Web Services Using Self-Adaptive Schema Hardening Algorithm

Vipul Patel, Radhesh Mohandas, and Alwyn Pais

Information Security Research Lab,
National Institute of Technology Karnataka, India
{vip04pat,radhesh,alwyn.pais}@gmail.com

Abstract. Web Services in production often evolve over time due to changes in business and security requirements. Often various Web Service standards such as WS-Security, WS-Trust, WS-Routing etc. are introduced or revoked. Such changes alter the structure of an input message accepted by web services. Message validation mechanism becomes in-effective if schemas in use are not updated in line with aforementioned changes. Also, Web Services become prone to different attack vectors if the schemas are loosely defined. Here, we present algorithms that help fine tune schemas by the process of iterative deduction. Also, our work helps to identify patterns of attack vectors that demarcate themselves from genuine messages. Our adaptive schema refining algorithm classifies logged requests into set of schema classes based on a measure of similarity. This classification of messages in to schema classes enables us to tighten the schemas to prevent bad requests or expand the schemas to accommodate newer requests.

Keywords: Schema Hardening, Schema Refining, Adaptive Algorithm, SOAP Message Validation, XSD Signature.

1 Introduction

A Web Service is used as a basic building block for the implementation of Service Oriented Architecture (SOA) based system. Growing importance of Web Services has enticed attackers. To safeguard web services from attackers, an extra layer of security is deployed at server side which checks for various vulnerabilities such as large payload size, well-formed message structure, XML Injections and many more. An easier way to hamper web services is by sending malformed messages that do not adhere to expected message structure. These kinds of messages either bring down the server or cause unintended operations on the server side. The best way to safeguard against such issues is to use message validation before the request is supplied to the business logic. A Simple Object Access Protocol (SOAP) message validator assures that messages adhere to the prescribed schema and discards messages if they appear to be malformed. Often such validators rely on Xml Schema Definitions (XSD) derived from Web Service Description Language (WSDL) document or hand coded by programmers.

The use of standards such as WS-Security, WS-Addressing, WS-Routing, WS-Trust, WS-SecureConversation etc. requires change in the structure of input SOAP messages as their use may introduce additional elements which were not present before. In the same way if the use of any such standard is revoked then the

D.C. Wyld et al. (Eds.): CNSA 2011, CCIS 196, pp. 383–392, 2011.

message structure can change significantly. Also, schemas used by validators may be loosely defined which may cause bad messages to pass through. These kinds of changes demand significant modifications to the schema being used for validation. Manual un-guided schema updates that completely rely on the skills of a programmer may make validator susceptible to attacks.

In this paper, we propose a solution to deduce groups of schema classes by the process of iteratively learning from the logged SOAP requests. This kind of classification provides clear demarcation between kinds of messages encountered by the server. Schema classes pertaining to good requests are used to tighten schemas and schema classes pertaining to bad requests are used to expand schemas. The act of fine tuning schemas based on the deduced schemas is called “Schema Hardening Process” that would increase the efficacy of the validation mechanism. It is adaptive in nature because schema classes are formed by learning message patterns.

Section 2 reviews related work pertaining to SOAP message schema validation and a mechanism of schema comparison. In section 3, we have discussed the overall architecture of our solution. Section 4 details an operation of our solution by describing working of each stage of an algorithm. Section 5 lists future work and then section 6 concludes the paper.

2 Related Work

Gruschka, N [4] has elucidated a need of message validation to thwart an attack on Web Services. They have shown the use of schema validation mechanism to counter the Denial of Service attack (DoS). DoS attack often relies on sending a message with large number of nested XML elements that would consume considerable server resources and keep it busy. Their work has shown a way to detect such oversized payload by validating them against hardened schema. Also, they have outlined a process of deriving schema from WSDL file. Web Services can also be compromised through XML signature wrapping attacks [5]. XML structure of SOAP messages affected by signature wrapping attacks differ significantly from the normal message structure. Our work leverages upon a similarity among XML schema to deduce schema classes. We have adapted algorithms discussed in [2] and [3] to calculate a measure of difference between candidate schemas. This schema comparison algorithm has its seed in the dynamic programming algorithm given in [1] that uses edit graph to determine a sequence of operations that would transform one schema tree in to another with minimum cost. Nierman and Jagadish [3] have described a notion of allowable sequence of edit operations by which an overall cost of comparison can be lowered significantly. The reduction in computation cost is achieved by means of graft costs and prune costs computations that they calculate for every sub tree of a schema tree. If a sub-tree of one schema tree is present in other schema tree then the cost of inserting/deleting whole sub-tree is taken into account and not just the cost of inserting/deleting every individual node of a sub-tree.

3 Overview of Adaptive Schema Hardening

All incoming SOAP request messages are intercepted by the message validator that validates them against schemas available in the schema repository. Initially, this