NGHIÊN CỨU BẢO MẬT DỊCH VỤ WEB, ỨNG DỤNG VÀO VIỆC TRUYỀN VÀ NHẬN DỮ LIỆU QUAN TRẮC MÔI TRƯỜNG

RESEARCHING WEB SERVICE SECURITY AND APPLY TO TRANSMIT AND RECEIVE ENVIRONMENTAL MONITORING DATA

Nguyễn Bá Nghiễn[1], Nguyễn Thái Cường[1], Vũ Tuấn Anh[2], Trịnh Trọng Chưởng[2]

TÓM TẮT

Trong bài báo này, chúng tôi trình bày kiến trúc và kỹ thuật bảo mật dịch vụ web thông qua việc bảo mật thông điệp SOAP XML bằng cách sử dụng chữ ký điện tử để đảm bảo tính toàn vẹn và mã hóa để bảo mật dữ liệu. Chúng tôi áp dụng chính sách bảo mật cho dịch vụ web vào việc truyền và nhận dữ liệu quan trắc môi trường từ các trạm quan trắc về máy chủ của các sở tài nguyên môi trường. Giải pháp của chúng tôi đưa ra giúp cho việc truyền và nhận dữ liệu được diễn ra một cách chính xác và an toàn mà các phương pháp trước đây chưa thực hiện được vì chúng chỉ truyền dữ liệu thô, không được mã hóa và xác thực nên dễ dàng bị tin tặc thay đổi nội dung trong khi truyền.

***Từ khóa**: Dịch vụ web, chữ ký điện tử, mã hóa, quan trắc môi trường, SOAP XML.*

ABSTRACT

In this paper, we present web service's architecture and technique for web service security by using digital signature and encryption to secure SOAP XML message. We apply this security policy to transmit and receive environmental monitoring data from monitoring station to resoure and environment department's server. Our proposal solution helps to transmit and receive data security and accuracy comparing to previous solutions beacause they transmit rough data which is not encrypted and signed so it is easy for hacker to change data when it is transmitting.

***Keywords**: Web service, digital signature, encrypt, environmental monitoring, SOAP XML.*

[1]Khoa Công nghệ thông tin, Trường Đại học Công nghiệp Hà Nội
[2]Viện Công nghệ HaUI, Trường Đại học Công nghiệp Hà Nội
*Email: chuonghtd@gmail.com
Ngày nhận bài: 07/01/2018
Ngày nhận bài sửa sau phản biện: 04/4/2018
Ngày chấp nhận đăng: 21/8/2018
Phản biện khoa học: TS. Phạm Văn Hà

1. GIỚI THIỆU

Khái niệm về dịch vụ web ra đời cuối năm 1990 và từ đó nó trở thành xương sống của ngành công nghiệp IT. Ngày nay, hầu hết các tổ chức kinh doanh đều dựa vào dịch vụ web để đạt được mục tiêu mong muốn của mình. Với sự khả chuyển và tùy biến mạnh của ngôn ngữ đánh dấu mở rộng (XML: Extensible Markup Language) nên nó trở thành ngôn ngữ phổ biến được sử dụng cho tất cả các dịch vụ web [1, 2]. Theo [3] XML web serices là mô hình thành công cho rất nhiều các ứng dụng web phức tạp. Giao diện của dịch vụ web được mô tả dựa vào XML và được gọi là Web Services Description Language (WSDL), việc trao đổi thông tin với dịch vụ web được thực hiện thông qua thông điệp XML SOAP (Simple Object Access Protocol). Do đó bảo mật dịch vụ web chính là đảm bảo tính toàn vẹn và bảo mật của thông điệp XML SOAP. Tổ chức World Wide Web Consortium (W3C) và Advancement of Structured Information Standards (OASIS) đã đề xuất một số tiêu chuẩn dùng để bảo mật dịch vụ web. Trong [2, 4] đã đưa ra một cái nhìn tổng quan về nâng cao hiệu suất xử lý các thông điệp SOAP cũng như tối ưu hóa bảo mật dịch vụ web và xử lý song song tài liệu XML. Trong nghiên cứu [5] đã cung cấp một cái nhìn tổng quát về các chuẩn bảo mật cho XML và dịch vụ web gần đây. Các chuẩn này cung cấp nền tảng đáp ứng các yêu cầu cơ bản về bảo mật như mã hóa, xác thực và đảm bảo tính toàn vẹn của dữ liệu cũng như các yêu cầu nâng cao như ủy quyền, liên kết danh tính.

Ở Việt Nam, theo yêu cầu của Bộ Tài nguyên và Môi trường thì tất các các khu công nghiệp xả thải ra môi trường với lưu lượng hơn 1000m^3 trên một ngày đêm đều phải lắp đặt các trạm quan trắc tự động các chỉ tiêu của nước thải, các nhà máy sản xuất thép và xi măng bắt buộc lắp trạm quan trắc tự động các chỉ tiêu của khí thải và truyền dữ liệu quan trắc về máy chủ của các Sở Tài nguyên và Môi trường. Để đảm bảo tính khách quan, Bộ cũng yêu cầu truyền dữ liệu lấy từ đầu ra của bộ cảm biến thông qua các bộ datalogger để truyền dữ liệu về các Sở Tài nguyên và Môi trường. Tuy nhiên, các bộ datalogger ở nước ta hiện nay chủ yếu sử dụng giao thức TCP/IP, UDP/IP hoặc FTP để truyền dữ liệu thô (không được mã hóa và xác thực) về các máy chủ. Điều này tạo ra lỗ hổng bảo mật cho các hacker dễ dàng bắt và thay đổi nội dung của các gói tin dẫn đến dữ liệu nhận được không chính xác. Trong bài báo này, chúng tôi đề xuất việc sử dụng dịch vụ web cùng với các chính sách bảo mật của dịch vụ web và SOAP XML để truyền và nhận dữ liệu từ các trạm quan trắc về máy chủ một cách chính xác và an toàn cao. Đề xuất của chúng tôi

đã được áp dụng thành công tại Sở Tài nguyên và Môi trường tỉnh Hải Dương.

2. CƠ SỞ LÝ THUYẾT

2.1. Dịch vụ web (web service)

2.1.1. Giới thiệu

Theo định nghĩa của W3C (World Wide Web Consortium), Web Service là một hệ thống phần mềm được thiết kế để hỗ trợ khả năng tương tác giữa các ứng dụng trên các máy tính khác nhau thông qua mạng Internet, giao diện chung và sự gắn kết của nó được mô tả bằng XML. Kiến trúc của dịch vụ web được minh họa trên hình 1.

Hình 1. Kiến trúc của dịch vụ web service [6]

Theo mô hình trên thì dịch vụ web sẽ gồm ba thành phần: Nhà cung cấp dịch vụ web (Web Service Provider), khách hàng sử dụng dịch vụ web (Web Service Consumer) và nhà môi giới dịch vụ web (Web Service Broker). Nhà cung cấp tạo ra các dịch vụ web và cung cấp cho các ứng dụng của khách hàng muốn sử dụng chúng. Khách hàng là các ứng dụng muốn sử dụng các chức năng mà dịch vụ web cung cấp. Môi giới dịch vụ web là ứng dụng cho phép các ứng dụng của khách hàng có thể tìm thấy các dịch vụ web đã được đăng ký. Ba thành phần này tương tác với nhau theo ba cơ chế như sau:

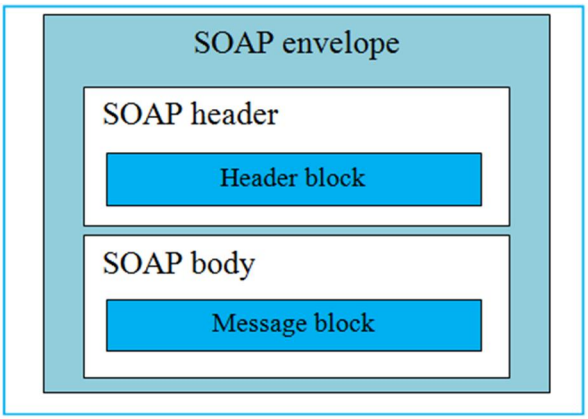
• Publish (phát hành): Nhà cung cấp dịch vụ web sử dụng giao diện chương trình của thành phần môi giới để đăng ký các thông tin liên quan đến dịch vụ web của mình như địa chỉ, các chức năng mà nó cung cấp để cho phép các ứng dụng của khách hàng có thể tìm và sử dụng đúng các chức năng mà nó cung cấp.

• Find (tìm kiếm): Ứng dụng của khách hàng dựa vào thông tin của các dịch vụ web đã đăng ký với Broker để tìm được dịch vụ web mong muốn.

• Bind (triệu gọi): Để sử dụng được dịch vụ thì cần phải triệu gọi nó. Trong thao tác này, ứng dụng của khách hàng khi thực thi sẽ gọi hoặc khởi tạo một luồng tương tác với dịch vụ dựa trên các thông tin trong mô tả dịch vụ mà nó thu được trước đó như: vị trí dịch vụ, cách liên lạc và tương tác với dịch vụ,…

2.1.2. SOAP (Simple Object Access Protocol)

SOAP là giao thức quan trọng trong Web service được xây dựng dựa trên XML, một giao thức truyền thông hay một định dạng để gửi thông điệp cho phép các ứng dụng trao đổi thông tin với nhau qua HTTP. Cấu trúc của thông điệp SOAP được minh họa trong hình 2 [6].



Hình 2. Cấu trúc của thông điệp SOAP

2.1.3. WSDL (Web services description language)

Web service không thể được sử dụng nếu nó không được tìm thấy. Chương trình của khách hàng muốn gọi một dịch vụ web thì nó cần phải biết dịch vụ web đó đang được lưu trữ ở đâu. Ngoài ra, chương trình của khách hàng cũng cần phải biết chức năng của dịch vụ web để nó gọi đúng dịch vụ web mong muốn. Công việc này được thực hiện thông qua sự hỗ trợ của WSDL (ngôn ngữ mô tả dịch vụ web). WSDL được xây dựng dựa trên XML và nó cung cấp cho chương trình ứng dụng của khách hàng biết chức năng của dịch vụ web cũng như vị trí của dịch vụ web. Như vậy, dựa vào tài liệu WSDL các chương trình ứng dụng của khách hàng xác định được vị trí và cách sử dụng dịch vụ web [6].

2.1.4. Universal Description, Discovery and Integration (UDDI)

UDDI là một chuẩn dựa trên XML định nghĩa một số thành phần cho phép các ứng dụng của khách hàng truy tìm và nhận những thông tin được yêu cầu khi sử dụng dịch vụ web. Một UDDI gồm có hai phần sau [6]:

• Phần đăng ký của tất cả các Web Service's metadata, bao gồm cả việc trỏ đến tài liệu WSDL mô tả dịch vụ.

• Phần thiết lập WSDL Port type định nghĩa cho các thao tác và tìm kiếm thông tin đăng ký.

2.1.5. Bảo mật dịch vụ web

Trong bài báo này, chúng tôi tập trung vào nghiên cứu việc truyền thông tin (yêu cầu và đáp ứng) với dịch vụ web thực hiện bằng thông điệp SOAP XML. Như vậy, việc bảo mật dịch vụ web chính là bảo mật thông điệp SOAP XML [7,8]. Các thông điệp SOAP yêu cầu và đáp ứng đều được ký với khóa của chứng thực (certificate) X509. Chữ ký phải tương thích với chuẩn về xử lý và cú pháp chữ ký XML và đáp ứng các yêu cầu sau:

- Chuẩn bảo mật dịch vụ web phiên bản 1.0 (WS - security 1.0) và chữ ký XML để thực hiện chữ ký số.

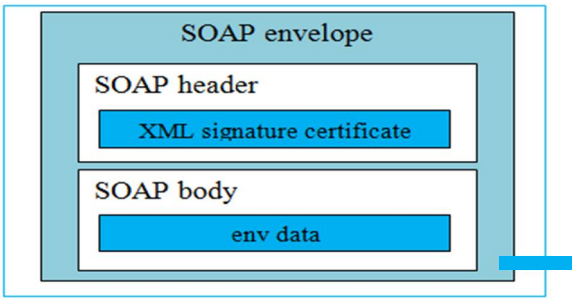
- Khóa bí mật của chứng thực X509 được sử dụng để thực hiện chữ ký điện tử trong phần thân (body) và đầu (header) của thông điệp SOAP được đính vào thành phần Binary Security Token theo định dạng của X509V3.

- Thuật toán "Exclusive C14N" được sử dụng để chuẩn hóa tài liệu XML nhằm thu được một chữ ký điện tử duy nhất.

- Để tính toán hash (digest) cho chữ ký của thông điệp SOAP nên sử dụng thuật toán SHA256.

- Sử dụng thuật toán RSA - SHA256 thực hiện chữ ký điện tử cho thông điệp SOAP.

Như vậy, cấu trúc của thông điệp SOAP có bảo mật sẽ như hình 3.

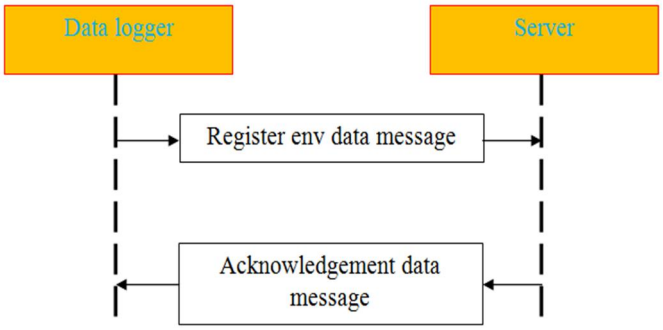


Hình 3. Cấu trúc của thông điệp SOAP có bảo mật

2.2. Đề xuất mô hình truyền dữ liệu từ trạm quan trắc về máy chủ của sở tài nguyên môi trường sử dụng dịch vụ web

2.2.1. Mô hình

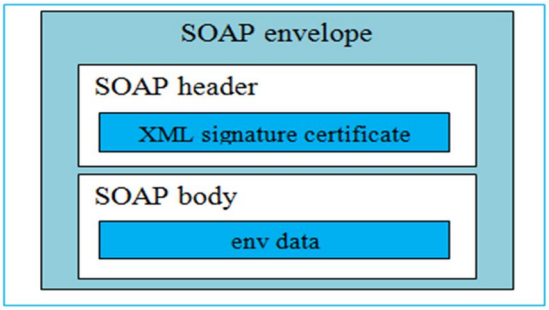
Mô hình đề xuất việc truyền và nhận dữ liệu giữa data logger và máy chủ của Sở Tài nguyên và Môi trường như hình 3. Bộ datalogger sẽ gửi các thông báo chứa thông tin về các thông số quan trắc môi trường của trạm đến máy chủ. Phía máy chủ sẽ kiểm tra tính toàn vẹn và cấu trúc của dữ liệu gửi. Nếu dữ liệu vượt qua việc kiểm tra thì sẽ được lưu vào trong cơ sở dữ liệu. Máy chủ cũng sẽ trả lời lại cho datalogger biết tình trạng nhận dữ liệu thành công hay thất bại.



Hình 4. Mô hình truyền và nhận dữ liệu

2.2.2. Cấu trúc của dữ liệu gửi và phản hồi

Tất cả dữ liệu gửi và phản hồi đều sử dụng SOAP và có cấu trúc như hình 5.



Hình 5. Cấu trúc chung của dữ liệu gửi và phản hồi

Định dạng XML của thông điệp yêu cầu gửi từ datalogger đến máy chủ của Sở Tài nguyên và Môi trường như sau:

<report:env>
<env:Data ***attributes***…/>
<env:Control>
values…
</env:Control>
</report:env>

Nội dung của các thuộc tính (attributes) và các giá trị được mô tả ở bảng 1.

Bảng 1. Mô tả chi tiết các thuộc tính và giá trị trong cấu trúc dữ liệu XML

Vùng dữ liệu	Mục	Mô tả	Tên thẻ XML
Data	1	Mã trạm quan trắc	code
	2	Tên trạm quan trắc	name
	3	Thời điểm quan trắc	date_time
	4	Chỉ số ph	ph
	5	Chỉ số cod	cod
	6	Chỉ số tss	tss
	7	Chỉ số màu	color
	8	Nhiệt độ	temp
	9	Lưu lượng tức thời	flow
	10	Tổng lượng nitơ hòa tan	nito
	11	bod	bod
Control	12	Chữ ký điện tử của trạm (station signature code)	ssic
	13	Mã bảo mật của trạm (station security code)	ssec

Trong đó:

- Mã trạm (code) gồm 3 chữ số từ 001 - 999.

- Tên trạm có độ dài tối đa 100 ký tự gồm các chữ cái, chữ số và dấu cách.

- Thời điểm quan trắc có định dạng dd-mm-yyyy Thh:mm:ss với dd là ngày 2 chữ số, mm là tháng 2 chữ số, yyyy là năm 4 chữ số, hh là giờ 2 chữ số, mm là phút 2 chữ số, ss là giây 2 chữ số.

- Các chỉ số quan trắc là số thập phân có 2 chữ số đằng sau dấu phảy.
- Chữ ký điện tử của trạm được thực hiện qua các bước sau:

Bước 1: Lựa chọn một số thông tin trong phần data làm dữ liệu được ký như sau: code| date_time|ph|cod|tss. Trong đó dấu '|' dùng để phân tách giữa các thành phần dữ liệu có mã ASCII là 124.

Bước 2: Dữ liệu được chọn ở bước 1 sẽ được ký điện tử bằng thuật toán SHA256withRAS với key và certificate dùng để ký toàn bộ dữ liệu gửi.

Bước 3: Dữ liệu sau khi được ký sẽ được mã hóa bằng thuật toán Base64.

- Mã bảo mật của trạm được xác định như sau:

Bước 1: Dữ liệu thu được ở bước 2 bên trên sẽ tiến hành các bước tiếp theo.

Bước 2: Sử dụng thuật toán SHA1 để tạo message digest.

Bước 3: Dữ liệu thu được ở bước 2 sẽ được mã hóa bằng giải thuật base16.

Bước 4: Chèn dấu '–' có mã ASCII 45 vào giữa các vị trí thứ 8 và 9; 16 và 17; 24 và 25; 32 và 33.

Ví dụ sau minh họa cấu trúc dữ liệu gửi:

```
<env:Report>
<env:Data code="100" name="Tan truong"
date_time="10-11-2016T22:05:00" ph="7.00"
cod="34.27" tss="12.35" color="12.74"
temp="20.12" flow="40.79" nito="0.00"
bod="0.00" />
<env:Control>
<env:ssic digest="SHA256" cipher="RSA2048"
encoding="base64">
Ca8sTbURReQjjgcy/znXBKjPOnZof3AxWK5WySpyMrUX
F0o7cz1BP6adQzktODKh2d8s
oAhn1R/S07lVDTa/6r9xTuI3NBH/+7YfYz/t92eb5Y6aNvL
m6tXfOdE3C94EQmT0SEEz
9rInGXXP1whIKYX7K0HgVrxjdxCFkZF8Lt12XbahhAzJ47
LcPxuBZZp6U6wJ2sWI5os3
KY9u/ZChzAUaCec7H56QwkMnu3U3Ftwi/YrxSzQZTmP
TpFYKXnYanrFaLDJm+1/yg+VQ
ntoByBM+HeDXigBK+SHaxx+Nd0sSmm1Im4v685BRVd
UId+4CobcnSQ3CBsjAhqmIrtWT
GQ==
</env:ssic>
<env:ssec digest="SHA1" encoding="base16">
03ec1d0e-6d9f77fb-1d798ccb-f4739666-a4069bc3
</env:ssec>
</env:Control>
</env:Report>
```

3. KẾT QUẢ NGHIÊN CỨU

Chúng tôi đã áp dụng mô hình truyền và nhận dữ liệu như đề xuất bên trên vào việc truyền dữ liệu từ các trạm quan trắc nước thải và khí thải tại các khu công nghiệp và nhà máy trên địa bàn tỉnh Hải Dương. Trên máy chủ của Sở Tài nguyên và Môi trường tỉnh Hải Dương chúng tôi xây dựng một dịch vụ web làm nhiệm vụ nhận và kiểm tra tính xác thực và toàn vẹn dữ liệu từ các trạm quan trắc gửi đến. Nếu dữ liệu chính xác sẽ tiến hành ghi vào cơ sở dữ liệu để thuận lợi cho việc truy vấn và hiển thị sau này. Phía datalogger sẽ cấu trúc dữ liệu gửi theo đúng khuôn dạng yêu cầu sau đó sử dụng dịch vụ web từ phía server để truyền dữ liệu. Hình 6, 7 minh họa kết quả mà chúng tôi nhận được từ các trạm quan trắc nước thải và khí thải của các nhà máy và các khu công nghiệp.

Hình 6. Minh họa kết quả truyền và nhận dữ liệu từ các trạm quan trắc tự động nước thải

Hình 7. Minh họa kết quả truyền và nhận dữ liệu từ các trạm quan trắc khí thải

4. KẾT LUẬN VÀ KHUYẾN NGHỊ

Trong bài báo này, chúng tôi đã đề xuất mô hình truyền và nhận dữ liệu từ các trạm quan trắc môi trường tự động về máy chủ của các Sở Tài nguyên và Môi trường của các tỉnh sử dụng dịch vụ web có bảo mật. Mô hình đề xuất của chúng tôi đơn giản, dễ dàng cho việc xử lý dữ liệu trên máy chủ mà không cần cài thêm bất kỳ phần mềm nào khác so với yêu cầu hiện tại của Bộ Tài nguyên và Môi trường là yêu cầu truyền từng file văn bản theo giao thức truyền tệp để

lưu trên máy chủ FTP. Do đó muốn hiển thị, truy vấn và phân tích dữ liệu cần phải viết bổ sung thêm phần mềm để đọc dữ liệu từ máy chủ FTP sau đó lưu vào cơ sở dữ liệu. Ngoài ra, mô hình đề xuất của chúng tôi cũng an toàn hơn so với các mô hình truyền dữ liệu hiện tại vì dữ liệu truyền trên mạng theo mô hình đề xuất được mã hóa và xác thực trong khi đó các mô hình truyền dữ liệu hiện tại là truyền dữ liệu thô (các file văn bản hoặc các thông điệp dạng text) nên rất dễ bị các tin tặc bắt và thay đổi nội dung của các gói tin dẫn đến dữ liệu nhận được bị sai so với dữ liệu gốc. Mô hình truyền và nhận dự liệu do chúng tôi đề xuất đã được áp dụng thành công cho việc truyền và nhận dữ liệu từ các trạm quan trắc môi trường nước thải và khí thái tại các khu công nghiệp và nhà máy trên địa bàn tỉnh Hải Dương. Dữ liệu mà chúng tôi truyền và nhận chính xác đúng như dữ liệu quan trắc tại nhà máy. Tuy nhiên, phương pháp mã hóa RSA 2048 mà chúng tôi sử dụng sẽ bị giải mã nhanh chóng khi máy tính lượng tử ra đời. Do đó, trong thời gian tới chúng tôi sẽ nghiên cứu các phương pháp mã hóa khác mạnh hơn để mã hóa và xác thực thông tin.

LỜI CẢM ƠN

Bài báo này được hoàn thành với sự trợ giúp kinh phí của đề tài KH&CN cấp tỉnh Hải Dương năm 2016, mã số: KTCN.29.TNMT.16.

TÀI LIỆU THAM KHẢO

[1]. Amazon web services: Overview of Security Processes, June 2013.

[2]. Joe M. Tekli, Ernesto Damiani, Richard Chbeir and Gabriele Gianini, 2012. *SOAP Processing Performance and Enhancement*. IEEE Transactions On Services Computing, Vol. 5, No. 3.

[3]. Nils Agne Nordbotten, 2009. *XML and Web Services Security Standards*. IEEE Communications Surveys & Tutorials, Vol. 11, No. 3.

[4]. Hongbing Wang, Joshua Zhexue Huang, Yuzhong Qu, Junyuan Xie, 2005. *Web services: Problems and Future Directions.*

[5]. Doug Tidwell, James Snell, Pavel Kulchenko, 2001. *Programming Web Services with SOAP*. First edition.

[6]. Heather Kreger, 2001. *Web Services conceptual architecture*. IBM Software Group.

[7]. Locktyukhin, Max; Farrel, Kathy (2010-03-31). *Improving the Performance of the Secure Hash Algorithm (SHA-1)*. Intel Software Knowledge Base (Intel), retrieved 2010-04-02.

[8]. IBM Corporation and Microsoft Corporation, 2002. *Security in Web Service World: A Proposed Architecture and Roadmap.* A Joint White Paper, http://schemas.xmlsoap.org/specs/ws-security/WSSecurity-Roadmap.htm