Contemporary Engineering Sciences, Vol. 7, 2014, no. 14, 691 - 697
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2014.4676

Design and Implementation of Electronic Authentication for Web Contents

Min Choi

Dept. of Information and Communication Engineering
Chungbuk National University
Cheongju, Chungbuk 361-763, Republic of Korea

Namgi Kim

[2] Dept. of Computer Science, Kyonggi University
Suwon, Kyonggi-do 443-760, Republic of Korea
Corresponding author

Abstract

In this paper, we focus on to authenticate web contents but web contents can be complex to authenticate. A website is made up of many parts consisting of static parts and dynamic parts. Dynamic parts are usually the data generated from database or computed at runtime. Static parts are usually the HTML/CSS3, javascript, and static images. The web contents especially dynamic parts have features certainly seem to have escalated in frequency. Thus, there have been occurred on online many problems such as injury to a person's honor or human right abuse. However, authenticating such a dynamic pages are very hard to achieve, because they changes very frequently. In order to resolve this problem, we present an electronic authentication approach for a snapshot of web contents. First, we take a feature extraction from images. Then, we make use of MD5 for data integrity for digital signature. Finally, we encrypt the data for security with the public key cryptosystem. This system enhances and verifies data integrity for a certain snapshot of web contents.

Keywords: Electornic authentication, message digest, hash, public key cryptography

1 Introduction

The motivation of this research is the necessity for taking a snapshot of a certain web page, such as a privacy policy agreement, contract between two parties, important data for submitting to courts, and so on. For this purpose, there are nothing for us to keep the electronic data for authentication. Even though if we take a snapshot for authenticating an web site content on a user's computer or smartphone, it is not officially approved. This is because digital/electronic data would be easily altered or modified. It is difficult to authenticate even if we print out the data onto papers. Thus there are still a lot of needs to authenticate online data. As an example, we consider a case of a newspaper's web site [2]. Assuming that the private policy of personal information which is collected through the newspaper web sites. The page explains the company's policy in detail. If there exist some problems between a web site user and the company, it is possible someone could disregard. The policy at the time of violation occurred or the related rules. However, company can easily alter or modify if there are problem and if the policy shown/announced on the web site is adverse for them. This is because the web server is in charge of the company and easily manageable. Then, the company may deny that policy as we have not announced such a policy on the web. These results are difficult to approve the existence and integrity of the existence of the policy on the web.

In order to resolve the authentication problem of electronics contents, we introduce how we can establish confidence in user identity electronically between person and person or between person and computers. Since the authentication and transaction take place across an open network such as the Internet, the authentication and transaction should take place over a controlled network. Before we go into more detail, we assume that information sent from a web server on behalf of a trusted content provider arrives with integrity at the client web browser. That means we are assuming there is no security threat on the communication channel or the channel is protected by a virtual private network or something others. Thus, we need not worry about a various security threats on Internet, we just focus on the electronic authentication technique only. First, we take a capture for snapshot image which is target to be authenticated. Then, we extract feature values from the snapshot image. For this, we convert the snapshot image, a finite sequence of data points, in terms of a sum of cosine functions oscillating at different frequencies. And we take the DC component from the converted result. Then, we make use of the feature values to generate 256bits MD5 hash data. Then, users upload and authenticate the data to our authentication server. From then on, in case there are any threats to the integrity of the data or in case someone has to approve the integrity of the snapshot, we can check the data integrity using our authentication system.

The rest of this paper is organized as follows: Section 1 describes introductions and related works. Section 2 explores design and implementation of our E-authentication system and presents the experimental results using our system.

Finally, we conclude and summary our work in Section 3.

2 Materials and Methods

This research focuses on design and implementation of the e-authentication system. To do so, we first introduce the e-authentication model and architecture. The e-authentication architecture, called e-auth, includes a trusted person and the E-Auth system. The E-Auth architecture shown in Fig. 1 includes various components: government, bank, E-Auth system, and user. Figure 1 shows the overall flow of our electronic authentication process in the E-Auth architecture. Before user start e-authentication, people who want to get the E-authentication should register to governments or banks with personal attributes in order to get the certification. The accepted forms of government-issued ID are state-issued photo ID card (i.e., driver's license) and passport. Bank which is shown in Fig. 1 has to ensure that the applicant's identity is one of the officially accepted, government-issued forms of identity. If all of the information is correct, the Bank endorses the request for the certification. After that, they can apply to a registration authority (our system) to become a subscriber of our service.

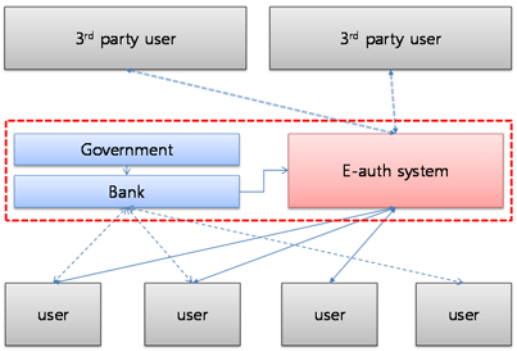

Figure 1 System Architecure

Each communication in Fig. 1 between government/bank and user or between e-auth system and user are either be RESTful open API interface or general HTTP web interface. The reason why we provide both interfaces is because we support web/mobile application and general PC applications, simultaneously. One of our key approaches in our e-auth system is that function of extracting feature value and generating MD5 hash value should be implemented onto e-auth system, not onto user client. This means that e-auth system provide the interface only, not the implementation detail, resulting in the perfect isolation between e-auth system and users. Thus, users recognize and access to e-auth system through RESTful open API interface. They don't have any accessible way directly, such as function call or library. As part of authentication, mechanisms

such as device identity could be used to identify or prevent possible authentication false positives. In this architecture, a public key and a related private key comprise a public key pair. The private key is stored on the token and is used by the Claimant to prove possession and control of the token. A Verifier, knowing the Claimant's public key through some credential (typically a public key certificate), can use an authentication protocol to verify the Claimant's identity, by proving that the Claimant has possession and control of the associated private key token.

Now we will discuss the procedure of processing a dynamic request, if server receives a request to the dynamic page from client, web server in kernel sends the cached static data to the client prior to constructing the complete dynamic page. First, web browser displays the static page to client. After that, user space web server sends the rest of the page, dynamic parts, to the client. Finally, the web browser integrates the dynamic parts into the static page displayed on the screen. So, this research supports to authenticate for a part of region which is specified by users. Then E-auth system take the histogram and feature extraction for a specific region of the snapshot, especially. This is necessary because E-authentication including dynamic data (for example, advertisements or something) will result in for the feature value extraction depending on the time when it will be extracted. We need to focus onto the not chaning part. The application redirects the user to the appropriate E-authorization endpoint, The following parameters are required: user certification, redirect_uri, and so on. The user logs into E-auth system with their credentials. The user is interacting with the authorization endpoint directly, so the application never sees the user's credentials. After successfully logging in, the user is asked to authorize the application. Note that if the user has already authorized the application, this step is skipped. Once E-auth system confirms that the client application is authorized, the end-user's Web browser is redirected to the callback URL specified by the redirect_uri parameter. Salesforce appends authorization information to the redirect URL with the following values: Since an image comprises hundreds or even thousands of 8x8 blocks of pixels, the following description of what happens to one 8x8 block in a microcosm of the

Figure 2 Sample Snapshots for E-Authentication

It should be noted that the pixel values of a black-and white image range from 0 to 255 in steps of 1. The values are between pure black 0 and pure white 255. Thus it can be seen how a photo, illustration, etc. can be accurately represented by these 256 shades of gray. Figure 3 is the histogram analysis result of Fig. 2.

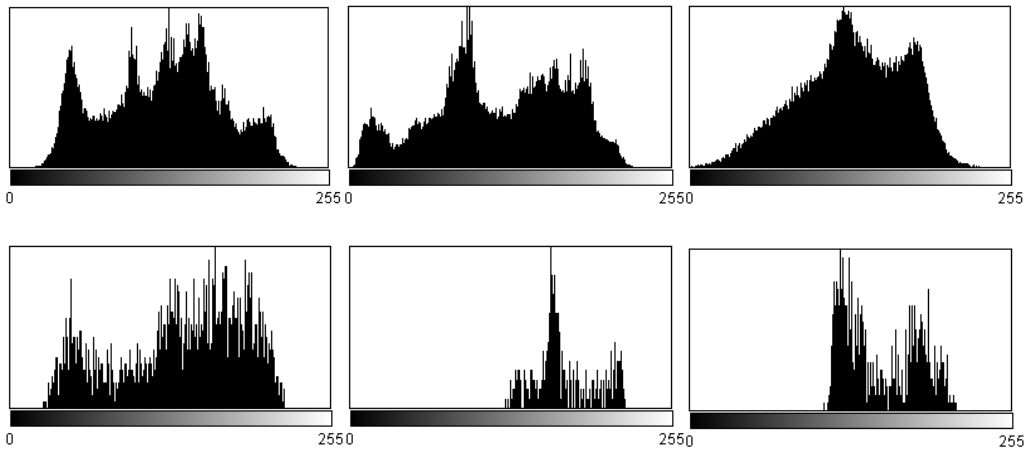

Figure 3 Histogram Analysis from the Sample Snapshots

In this research we utilize such histogram analysis data for feature vectors of the snapshot. Thus, both feature values from this section and feature values from this Section constitutes the feature values of a certain snapshot image. Moreover, here a user has to specifiy a region of interest, directly. Then, our E-auth system will analysis and generate MD5 hash and the digested results will be attached to the original snapshot image.

3 Conclusion

As digital information becomes more pervasive, the threat to the integrity of the data increases. The proliferation of electronic services, e.g., Internet shopping, Internet banking, e-learning, has changed in many aspects of daily lives. Clearly, there are a lot of demands for an efficient authentication system that will guarantee the integrity of the electronic content from a trusted source to a client, especially when it is distributed and aggregated through third parties. Web content can be complex to authenticate. A website is made up of many parts (html, stylesheets, javascripts etc.). The main focus in this paper is authentication of web content. The web contents have features certainly seem to have escalated in frequency. Web data cover a variety of platforms, metadata and file types, so electronic authentication of web contents must take care to choose a method that collects the content and data required for authentication. Just time stamp and a hash value are not enough for authenticating web contents. In order to resolve the authentication problem of electronics contents, we present a authentication approach for a snapshot of web contents on PC or mobile devices. This technology leverages a secure authentication by extraction of the feature values from image, then we apply MD5 message digest to the feature values.

Acknowledgements. This work is jointly supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0008105).

References

[1] D.H. 1. M. T. Goodrich, D. Nguyen, Efficient Verification of Web-Content Searching Through Authenticated Web Crawlers International Conference on VLDB

[2]. Arun C. Murthy, Chris Douglas, Mahadev Konar, Owen O'Malley, Sanjay Radia, Sharad Agarwal, Vinod K V, "Architecture of Next Generation Apache Hadoop MapReduce Framework", Technical Reports, 2013

[3]. Processing and Loading Data from Amazon S3 to the Vertica Analytic Database, Amazon Web Service, White Paper. 2013

[4]. Amazon Elastic MapReduce Developer Guide, Amazon Web Service, 2009.

[5]. Getting Started with Amazon Elastic MapReduce, Amazon Web Service, March 2009.

[6]. Md. Syeful Islam, Md. Rezaur Rahman, Anupam Roy, Md. Imdadul Islam and M. R. Amin, "Performance Evaluation of Finite Queue Switching Under Two-Dimensional M/G/1(m) Traffic," Journal of Information Processing Systems, vol. 7, no. 4, pp. 679~690, 2011

[7]. Rong Pan, Guandong Xu, Bin Fu, Peter Dolog, Zhihai Wang, Martin Leginus, "Improving Recommendations by the Clustering of Tag Neighbours", Journal of Convergence,.vol. 3, no. 1, pp. 13~20, Mar. 2012.

[8]. Haibo Zhao and Prashant Doshi, "Towards Automated RESTful Web Service Composition", Inteetrnational Conference on Web Services (ICWS), 2009.

[9]. Xia Zhao, Enjie Liu, Gordon J. Clapworthy, Na Ye, Yueming Lu, "RESTful Web Service Composition: Extracting a Process Model from Linear Logic Theorem Proving", IEEE international conference on Next Generation Web Service Practice(NWeSP), 2011.

[10]. Zheng Li and Liam O'Brien, "Towards Effort Estimation for Web Service Compositions using Classification Matrix", 2010

[11]. Cesare Pautasso, Olaf Zimmermann, Frank Leymann, "RESTful Web Services vs. Big Web Services : Making the Right Architectural Decision", International Conference on WWW 2008.

[12]. Rosa Alarcon, Erik Wilde, and Jesus Bellido, "Hypermedia-Driven RESTful Service Composition"

[13]. Cesare Pautasso, "RESTful web service composition with BPEL for REST," Data and Knowledge Engineering, vol. 68, no. 9, pp. 851-866, September 2009.

[14]. J. Rao and X. Su, "A survey of automated web service composition methods," In Semantic Web Services and Web Process Composition, pages 43-54, 2004.

[15]. J. Dean and S. Ghemawa, "MapReduce: Simplified Data Processing on Large Clusters", 6th USENIX Symposium on Operating System Design and Implementation, 2004.

[16]. Min Choi, Joung-Hyuk Park, Young-Sik Jeong, Mobile cloud computing framework for a pervasive and ubiquitous environment, The Journal of Supercomputing, Volume 64, Issue 2, pp 331-356, May 2013.

[17]. The Internet of Things: In action, The Next Web, http:// thenextweb.com/insider/2013/05/19/the-internet-of-things-in-action/
[18]. M. Yoon, Y.K. Kim, and J. W. Jang, “An Energy-efficient Routing Protocol using Message Success Rate in Wireless Sensor Networks”, Journal of Convergence, Vol. 4, No. 1, Mar. 2013.
[19]. F. Ozgur Catak, M. Erdai Balaban, “CloudSVM: Training an SVM Classifier in Cloud Computing Systems”, Pervasive Computing and the Networked World, pp. 57-68, 2013.
[20]. H. Yang, A. Dasdan, R. Hsiao, D. Parker, “Map-reduce-merge: simplified relational data processing on large clusters”, ACM SIGMOD international conference on Management of data, pp. 1029-1040, 2007.
[21] M. de Kruijf and K. Sankaralingam, “MapReduce for the Cell B.E. Architecture”, Vertical Research Group. Department of Computer Sciences, University of Wisconsin-Madison. 2010.
[22]. Hadoop, http://hadoop.apache.org/
[23]. Amazon Web Service, http://aws.amazon.com
[24]. IBM Smart Cloud, http://www.ibm.com/cloud-computing/us/en/
[25]. Komal Mahajan, Ansuyia Makroo and Deepak Dahiya, "Round Robin with Server Affinity: A VM Load Balancing Algorithm for Cloud Based Infrastructure," Journal of Information Processing Systems, vol. 9, no. 3, pp. 379~394, 2013.

Received: May 1, 2014