CONTEXT-AWARE ANALYSIS FOR ADAPTIVE UNIFIED AUTHENTICATION PLATFORM

Khairul Azmi Abu Bakar[1] and Galoh Rashidah Haron[2]

[1]Information Security Lab, MIMOS Berhad, Malaysia, khairul.azmi@gmail.com
[2]Information Security Lab, MIMOS Berhad, Malaysia, rashidah@mimos.my

ABSTRACT. Context-aware user authentication has become a necessary tool to serve as an additional security perimeter to protect online identities. By learning about user's online behavior over time, authentication system are able to establish user behavior profile. Any substantial deviation on log-in context from the profile would indicate a high risk login attempt. Such adaptive authentication system has been developed in production server where user login records for more than six months have been collected. In this paper, the analysis of the user login context is presented.

Keywords: adaptive authentication, web application, context-aware authentication, risk-based authentication system

INTRODUCTION

User authentication is the most basic form of security. User authentication provides the ability to prove identity of the user before the access to the information and resources the user entitled to is granted. Proving identity involves presenting one or more credentials from three possible factors: something you know (passwords, PINs, etc.), something you own (driver's license, token, corporate badge, etc.) and something you are (biometric: face, fingerprint, voice, retina, etc.).

Relying on just the credential to protect system from attacker is definitely not a wise decision. For example, username and password can be broken using several techniques such as brute force attacks and shoulder surfing (Raza et al., 2012). As a result, additional security parameter is required. Context-aware user authentication has become a necessary tool to serve as an additional security perimeter to protect online identities. In order to do so, authentication system must be able to understand the context of a login attempt, analyze that context to determine risk and adapt requirements accordingly.

Examples of such systems are (Shi, 2011) and (Rocha, 2011) but they are designed for use in mobile devices. We built authentication system for accessing to web applications and put it on production server for couple of months. The system had been collecting and storing users' login information into a database. In this paper, we present the analysis result of the context of the login attempts.

BACKGROUND

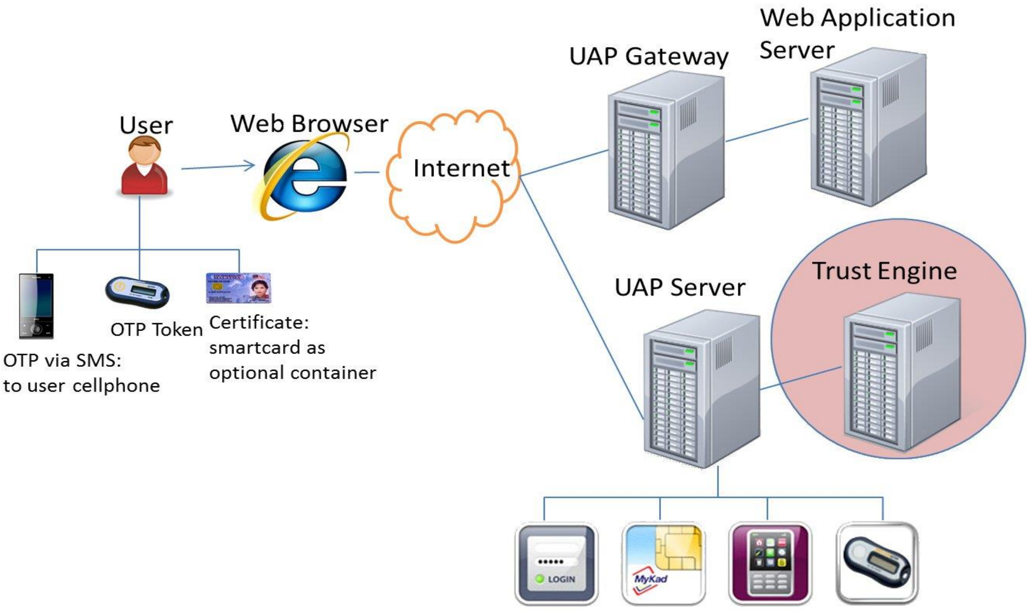

Figure 1. Adaptive UAP Architecture Diagram

In MIMOS Berhad, we have developed an authentication system called Unified Authentication Platform (UAP). UAP is a centralized multi-factor authentication system with web-based single sign-on (SSO) capability to manage user authentication profiles. UAP is derived from Shibboleth (Shibboleth, 2015) which is a standard based, open source package for web single sign-on across or within organizational boundaries. Users can choose from a list of authentication methods to get authenticated and be allowed to access multiple applications without having to go through the same authentication process again. The overall architecture of UAP is depicted in Figure 1 (Haron et al., 2012).

For the new generation of UAP, called adaptive UAP, we introduce an additional component called Trust Engine that incorporates adaptive control based on security risk and level of assurance (Abu Bakar et al., 2013). Trust Engine receives requests from UAP Server that contains login information as shown in Table 1. To make informed authentication decision, Trust Engine takes into account user common attribute from the behavior profiles which had been previously analyzed and stored in a dedicated database table data_log (Abu Bakar et al., 2014).

Table 1. User Login Information

No	Entry	Description
1.	uuid	User unique identity
2.	time_login	Data and time of the login
3.	browser_osname	User’s browser and OS information
4.	ip_int	User’s terminal IP Address
5.	sp_id	UAP Gateway Server ID string
6.	auth_method	Authentication method ID
7.	Tr	Application required trust level

COMMON CONTEXT

In this paper, four attribute factors mainly geographical location, time login, browser/operating system and accessed application are used for analysis. For each attribute factor, the system continuously analyze user login records to determine the common context. Common context should meet the following conditions (1) number of records for the last 14 days is more than 10 and (2) frequency of occurrence of any particular context is more than 30% of the overall records.

If there exist common profile within the attribute factor, the system compares each common context with the current login context. Every time the user logins under different login context, the number of that event is recorded.

EXPERIMENT RESULT

We collected login information from our production server from 6 May 2014 until 15 January 2015. Total number of 171,045 login information from 1244 unique users have been recorded during those 254 days period.

Geographical Location

The information about user geographical location is extracted from the IP Address of the user terminal. We use solution from a third party company ip2location [ip2location, 2015] which provides database records that contain geographical information such as name of the city, region and country of origin for IP Addresses. Special IP Addresses (10.0.0.0 to 10.255.255.255, 172.16.0.0 to 172.31.255.255, or 192.168.0 0 to 192.168.255.255) are originated from internal network and are categorized as private IP Addresses.

From the total number of records collected, 67.7% (115,816) have IP Address information. 97.8% (113,254) of those login records that contains IP Address information are from internal network. From the other 2,562 login records that comes from external network, 2,515 (98.1%) are from Malaysia where city of Kuala Lumpur is the most originating login access location with 2,116 records (84.1%). The remaining login records are from USA (37), Sweden (3), Thailand (3), Netherland (2) and Philippines (2).

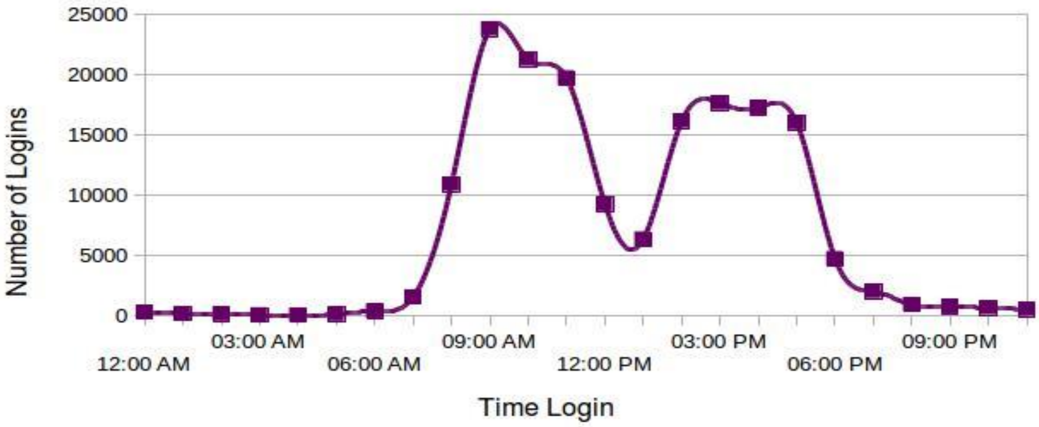

Figure 2. Number of Login with Respect to Time Hour

Time Login

Time login entries stored in table data_log are the time of the server machine when the login requests were received from UAP server. The graph in Figure 2 shows that the number of users login increases from 8 am onwards. After 6:00 pm, the number of login starts to decrease. This is normal since the standard working hours is between 8:30 am to 5:30 pm. In the afternoon, the number of login decreases for around 2 hours before increases back. It is a typical time for users to have their lunch break and stop operation.

Type of Browser and Operating System

Adaptive UAP is able to extract information such as type of browser and operating system running at the user terminal the agent string header send by the browser. There are ways to modify the string such as by using browser extensions (User Agent Switcher (UserAgentSwitcher, 2015) and User Agent Selector (UserAgentSelector, 2015)). However, in this case study, we assume that all user agent strings received by the system are authentic and unmodified.

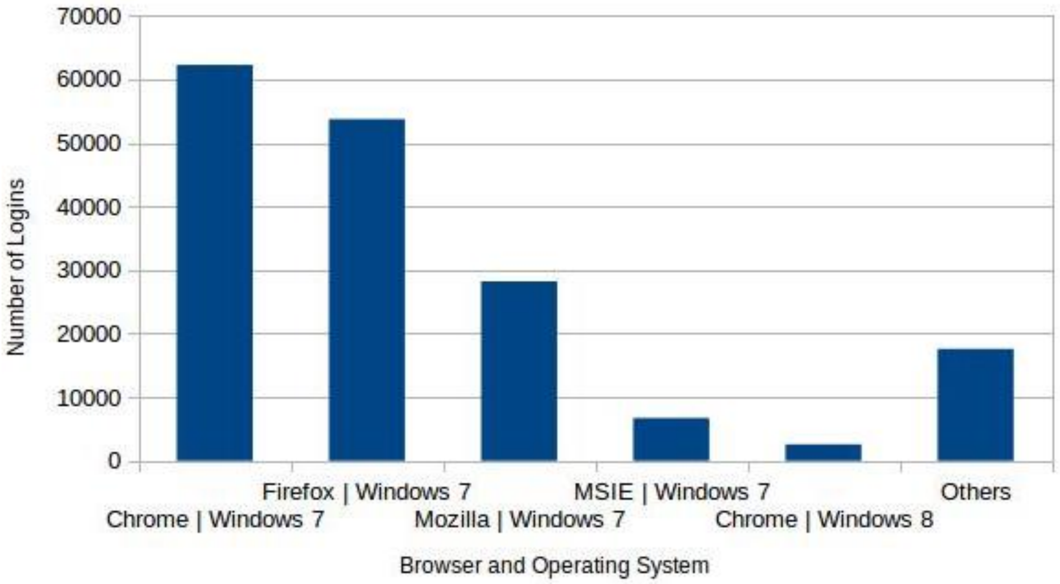

Figure 4. Number of Logins with Respect to paired Browser and OS

Trust Engine regards both entries of browser and operating system as one login context. If one of them is different, Trust Engine consider it as a different entry. Figure 4 shows the number of the paired value of browser and operating system derived from the collected login records. The top four most favorite operating system is all Windows 7. In those top list, the most popular browser is Chrome (36.4%), followed by Firefox (31.4%), Mozilla (16.5%) and Internet Explorer (3.9%). The fifth place is browser Chrome with Windows 8 operating system which is 1.5% from the total number of collected login records.

Application

Six UAP gateway servers have been used with only one of them was intended for high trust application. Table 2 shows the list of all UAP gateway servers used.There could be more than one application associated with one UAP gateway. If a user accesses two different applications which are both behind one UAP gateway, Trust Engine will get the same UAP gateway information for both login attempts.

Table 2. List of all UAP Gateway Servers

No	UAP Gateway Server ID
1.	https://uap.mimos.my/shibboleth
2.	https://sp.essweb.mimos.my/shibboleth
3.	https://bmyhdw.mimos.com/8647196938
4.	https://ipms-uap-gw-prod/vfdst35hj3120
5.	https://misso.mimos.my/shibboleth
6.	https://HT-MIESS1.mimos.my/shibboleth

Based on the collected login records, we found that majority (87%) of the logins are for accessing UAP gateway labelled as https://sp.essweb.mimos.my/shibboleth. The second most popular UAP gateway is https://HT-MIESS1.mimos.my/shibboleth which is assigned for high trust application. High trust applications are used to view high confidential documents. Examples of high trust applications are e-payslip to view the salary details online and e-PCB that shows employer's monthly income tax payment to the government tax return agency.Next, we analyze the number of login event where the users login from different context from the common attribute. As shown in Figure 5, type of browser contributes the highest number of events followed by application, operating system, time login and location. The last column labelled 'none' represents the login events where the users login under common context. We can conclude that users as individual have more tendency to use different type of browser compared to other factors. Location is the least factor. One main reason is because most of the users are the employers of the company and most of the applications can only be accessed from the company private network. Based on this result, location is the most influential factor and should be given the highest weightage in calculating the attribute factor score.

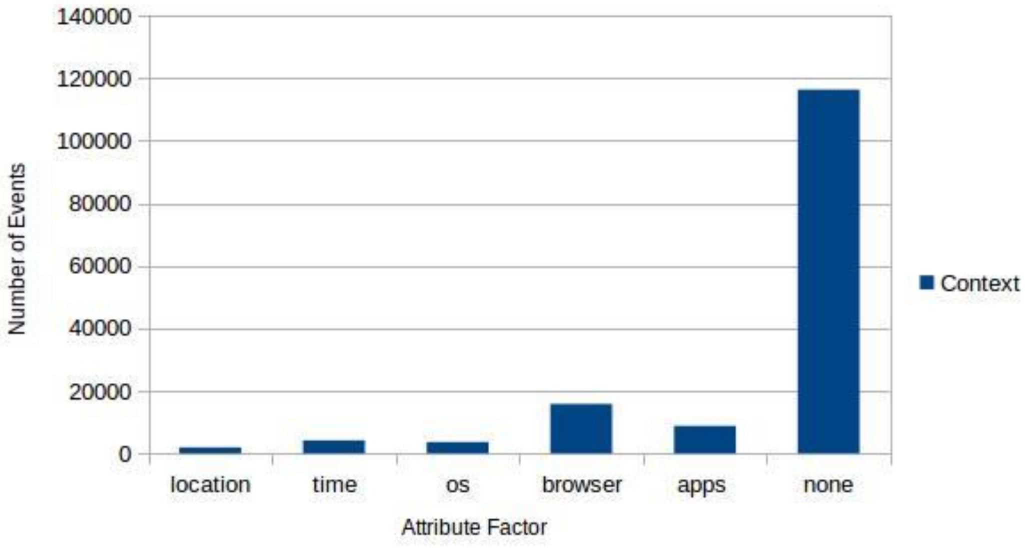

Figure 5. Total Number of Events for every Attribute Factor

CONCLUSION

Adaptive authentication is an additional security layer that uses risk factor analysis to make authentication decision. Users' past login records are used to form user behavior profile. Context information in login records such as user geographical location and time login are important parameters to reflect user common behavior profile. If the user logins from a different environment from the established behavior profile, adaptive authentication calculates the risk associated with the deviation and may request the user to present additional authentication method to be authenticated. In order to get the influential order of attribute factors, we did analysis on actual users login records captured from the production environment. The result suggests that location is the most influential factor, followed by login time, operating system, targeted application and browser type.

ACKNOWLEDGMENTS

We acknowledge the support provided by Ministry of Science, Technology and Innovation (MOSTI) in funding the MIMOS Unified Authentication Platform (UAP) project through the Tenth Malaysia Plan (10MP). The completion of the project allows the delivery of a centralized authentication infrastructural platform for web applications.

REFERENCES

Abu Bakar, K. A., & Haron, G. R. (2013, June). Adaptive authentication: Issues and challenges. In *World Congress on Computer and Information Technology (WCCIT)*, 1–6.

Abu Bakar, K. A., & Haron, G. R. (2014, August). Adaptive authentication based on analysis of user behavior. In *Science and Information Conference (SAI)*, 601–606.

Haron, G. R., Maniam, D., Sadavisam, V. & Loon, W. H. (2012). Re-engineering of web reverse proxy with shibboleth authentication. In *The 7th International Conference for Internet Technology and Secured Transactions (ICITST-2012)*, 325–330.

Ip2location. (2015). ip2location home page. http://www.ip2location.com. [Retrieved April 2015].

Raza, M., Iqbal, M., Sharif, M. & Haider, W. (2012), A Survey of Password Attacks and Comparative Analysis on Methods for Secure Authentication, *World Applied Sciences Journal*, *19(4)*, 439-444.

Rocha, Cristiano C., Lima, Joao Carlos D., Dantas, M. A. R., & Augustin, Iara. (2011, June). A2best: An adaptive authentication service based on mobile user's behavior and spatio-temporal context. In *IEEE Symposium on Computer and Communication (ISCC)*, 771–774.

Shi, Elaine, Niu, Yan, Jakobsson, Markus, & Chow, Richard. (2011). Implicit authentication through learning user behavior. *Information Security*, Lecture Notes in Computer Science, 99–113.

Shibboleth. (2015). Shibboleth documentation. https://wiki.shibboleth.net. [Retrieved April 2015].

UserAgentSelector. (2015) https://chromeuseragentselector.wordpress.com/. [Retrieved 30 January 2015].

UserAgentSwitcher. (2015). http://chrispederick.com/work/user-agent-switcher/. [Retrieved 30 January 2015].