Authentication Based on Context and Behavioral Biometrics

While authentication methods based on conventional passwords are still the most widely used, they offer a low level of security and are vulnerable to multiple attacks used for identity theft. Security strategies based on the analysis of context information and behavioral biometrics enhance the security provided by passwords.

A growing number of security services take into account the specific context in which operations takes place. Known as context-aware security services, they make use of all the additional information that can contribute to, among other aspects, improving decision making in authentication and authorization.

In the case of authentication, the context analysis is based on the fact that each user has a set of reasonably stable habits. A user connects at certain times from certain access points (from home or work) and uses a limited number of devices (computers and smartphones). When these user habits are known, anomalies can be detected that may indicate an attempt to steal the legitimate user's identity.

In certain environments, such as at work, the risk of user habits being impersonated is greater. The attacker who knows the user's password can replicate the user's habits and use the device under the legitimate user's identity. Safelayer's approach provides an additional layer of protection based on behavioral biometrics that is capable of detecting attempts to impersonate habits and the illegitimate use of devices.

Context and How It Is Captured

The user's authentication context is determined by certain characteristics and data of the environment in which the authentication process takes place, including the device,

connection network, location, time range and other data on user behavior. If over various authentications the context remains constant or is subject only to slight variations, the probability that the authenticated user is the legitimate owner of the credentials is greater than if the opposite occurs, i.e., the context changes noticeably over different authentication processes. The distance between contexts is determined in the context analysis phase and provides, as a result, the authentication's level of risk. Basically, a greater distance between contexts means there is a higher risk that an attempt is being made to fraudulently use the identity of the legitimate user.

The capture and analysis of the context factors does not affect the user experience. Therefore, the reduction in risk and improvement in authentication security is achieved almost totally transparent.

The client-side capture is done transparently using code that the authentication server securely sends to the user's browser or application. This code is used to capture device information such as the types and versions of the browser and operating system, monitor characteristics, device language and time-zone, sources and plug-ins installed, etc. These parameters are securely sent to the server, which compiles a fingerprint of the user's device. The device's fingerprint can be compared to the fingerprints of other devices the user has used in other authentications. Device fingerprints are compared using a smart algorithm that can detect the degree of coincidence based on the data. For instance, if only the browser version has changed, two different fingerprints can correspond to the same device. Devices are, however, considered different if the operating system is different.

The combination of the device fingerprint analysis and the use of one-time cookies constitute a complex device identifier as specified by the US's Federal Financial Institutions Examination Council (FFIEC). Furthermore, on the client side, network information is also captured in the form of the IP address and, in the future, geolocation (latitude and longitude). This last factor depends on the device type, i.e., for desktops, the Wi-Fi network can be used, whereas for mobile environments, the GPS or GSM systems can be used.

On the client side, authentication date and time are captured, the network information is recaptured and the geolocation is obtained from the IP address. Analyzing the geographic location can flag risk situations when consecutive authentications are attempted from locations that are geographically too far apart. The client can also be used to capture biometric information on the user being authenticated. Specifically, for username and password-based authentications, keystroke dynamics is captured and is used as a distinctive behavioral biometric trait. So, while context parameters are used for identifying the device, keystroke dynamics are used to improve user identification.

Using the information compiled in the context analysis is optional. It is up to operator of the authentication solution to decide which parameters are more important and should be used as authentication factors.

Using Keystroke Dynamics to Enhance Authentication

Just as we generate a personal and distinctive stroke when we handwrite, a series of unique characteristics can also be detected when we type. This characteristic is a biometric trait known as keystroke dynamics. It can be used to strengthen authentication processes.

Keystroke dynamics is a biometric trait related to user behavior.

One of the advantages of the keystroke dynamic analysis system compared to other biometric systems is that it is a non-intrusive method that can be used in a way that is 100% transparent to the users. Furthermore, parameters are captured via the keyboard, which means no additional hardware sensor is required. The capture is performed in the applications themselves (e.g., using JavaScript code interpretable by all browsers), which means no specific software is required. The keystroke analysis system can be enabled while users continue using their credentials as they always have.

The advantage of using a factor related to behavioral biometrics is that it offers a high rate of identification with minimal intrusion in the user experience. This characteristic allows detecting potential identity-theft attacks that take place with or without the legitimate user's knowledge. In other words, either because the attacker manages to steal the credential or because the legitimate user gives it to a third-party. In both cases, keystroke dynamics detects the improper use of the credential.

The keystroke dynamics analysis can be very useful for detecting an attacker using the credentials of a legitimate user in corporate environments, where it is very easy to reproduce other authentication factors such as the connection time and location, and where it is even relatively feasible to obtain the credential and the device of another member of staff.

TrustedX Adaptive Authentication can recognize if the keystroke dynamics captured and analyzed in each authentication process it coordinates corresponds to the legitimate user. Safelayer has designed and implemented its own analysis system.

The main parameters that characterize the keystroke dynamics of each user and that can be used to generate the user's biometric pattern are 1) the time a key is pressed (hold time or dwell time), 2) the time elapsing between the pressing of two keys (latency) and 3) the time elapsing from the release of one key and the pressing of the next (flight time). These time parameters are complemented by others that help to define the user's typing pattern with greater accuracy. For example, the system captures i) whether upper case characters are performed with the SHIFT key or with the Caps Lock key, ii) whether numerals are entered with the numeric keypad, iii) the number of times that the DELETE or BACKSPACE key is pressed to determine whether users usually make mistakes when entering their credentials, iv) whether the TAB key or the mouse is used for moving between fields in a form and v) whether the ENTER key or the mouse is used for sending credentials.

TrustedX's keystroke dynamics analysis component does not need to know the code of the keys pressed, which means using it does not compromise the confidentiality of the passwords.

System learning

To perform an authentication context analysis, the system must have a prior knowledge base on the typical behavior of the users. This allows the system to determine if the authentication is being performed under reasonable conditions or if it is an anomalous situation that indicates a possible case of identity theft. To build this knowledge base, the authentication system must define the conditions of the learning phase.

As in any biometric system, the users must pass an initial registration phase (also known as enrollment or training) so that the analysis of their keystroke dynamics can be performed subsequently. In this training phase, users enter their credentials multiple times so that the system can build their biometric patterns. This training process can be explicit, in which users repeatedly enter their credentials in an interface expressly for training, or transparent, in which the system stores samples of the user's keystroke dynamics over consecutive authentications.

After the initial training phase, the user's pattern can be updated with new samples captured during authentication processes as users may continue to change their keystroke dynamics as they get used to entering their credentials.

The samples entered by the user must be similar enough to each other for the biometric pattern to have the minimum quality for guaranteeing the correct operation of the analysis system. In an explicit training process, the user can see the quality of the pattern being generated as more samples are added. This means that the user can discard samples that worsen the quality of the training and even entire typing patterns. The keystroke analysis system stores the typing pattern of each user. In each user access, the captured typing sample and the typing pattern of the user that owns the credentials entered are compared. If the degree of coincidence between the sample and the pattern is very high, the guarantee that the user being authenticated is the owner of the credentials is increased. However, a low level of coincidence between the sample and the pattern might mean that the user entering the credentials is not the owner of them.

While keystroke dynamics are captured when the credentials are entered, the recognition algorithm for this factor comes into operation after the credentials have been validated to strengthen the authentication. In other words, the keystroke dynamics analysis is used to corroborate the identity of the user and verify that the user is not an attacker. It is not used to identify the user.

As users can access from multiple devices, their keystroke pattern may be different in each device, depending on the characteristics of each keyboard. Thus, users should be allowed to carry out the training phase in different devices so different patterns can be taken into account.

The keystroke dynamics analysis component acts independently of the device in which the pattern was generated. This means that if the user types in the same way in different devices, there is probably no need for more than one training session.

For smartphones and tablets with touchscreens, the keystroke dynamic capture method is exactly the same as for devices with desktop keyboards. However, the error rates of the analysis and registration processes are higher because 1) users usually use these devices in different positions, even when moving, and 2) typing is more irregular and imprecise than on a desktop keyboard. For this reason, other biometric traits that are

more suited to mobile devices while being just as effective in terms of identification, such as facial recognition and movement analysis, are being investigated.

Safelayer's biometric system has a hit rate of over 95%. False negatives are reduced by making use of other context factors in the analysis.

Context and Authentication Factors

The context analysis provides authentication factors in addition to the *something you know* factor, i.e., the user's password. Each of these factors can be seen as an additional layer of security that strengthens the password and enhances the protection of the system:

- Something you have, i.e., the analysis of the user's device, in which the probability that the device being analyzed is in possession and under the control of the legitimate user is assessed.
- Something you do, i.e., the analysis of the usual location and connection time-range of the legitimate user.
- Something you are, i.e., the biometric analysis of the legitimate user's keystroke dynamics.

In addition to the individual analysis of each of the above factors, a cross analysis can be carried out by correlating several factors and comparing it to the authentication activity history of the user. These correlations are different combinations of device, location and time range.

TrustedX Adaptive Authentication analyzes the context in which the authentication takes place and assesses the risk that an attempt of identity theft may be taking place.

This analysis can be carried out without affecting the normal user experience. It can enrich any authentication process in which a username and a password are used. It is therefore especially useful in environments where it may not be possible to use stronger methods, for example, for reasons of usability (as usually occurs in online shopping scenarios) or because the cost of implementing stronger methods (such as a OTPs hardware) is prohibitive.

So, the authentication context analysis detects anomalies in authentication processes, increases the level of security provided by passwords, reduces the risk of identity theft and provides a greater degree of trust in the protection of resources.

The system stores detailed information on all the context factors of the authentication processes. Furthermore, when an anomalous situation is detected, the system generates an event that is logged in the auditing system. This event can be converted into an alert or configured to activate an additional authentication factor for verifying that the user is the owner of the credentials.