User Identification Based on Standard Input Devices Usage

Peter KRÁTKY*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
kratky@fiit.stuba.sk

Every person is unique in the way he/she uses a computer, an operating system, programs and even input devices such a keyboard or a computer mouse. Patterns produced by usage of standard input devices could be utilized to identify an unauthorized user. Especially, adaptive and personalized systems accessed by multiple users might benefit from this feature as tailoring the content heavily depends on previous activity of the user. In our research, we focus on patterns in browsing the web, a very common activity nowadays. The goal of our work is to design a method to identify persons based solely on their behaviour rather than machine or browser they use to access websites.

Construction of a user model representing input device usage patterns has been a problem discussed by many researchers. The common protection of a user account by password is enriched with mechanism of patterns comparisons when accessing the account. Apart from comprehensive studies of keyboard usage dynamics, numerous papers focus on computer mouse. One of the first works [1] has proven the possibility of identification by characteristics such as distance, duration and angle of the mouse movement path. Comprehensive list of spatial and temporal characteristics of mouse movement strokes (curvature, acceleration, angle velocity, etc.) have been published and studied in [2]. Analysis of changeability of characteristics according to the environment (display resolution, mouse sensitivity) was described in [3]. We focus on identification which is harder task, but the previous research shows it is a possible one.

The process of user identification consists of three stages – (I.) acquisition of the data while browsing, (II.) construction of a user model and (III.) the identification itself. We designed a system consisting of three modules corresponding to these stages.

The first part of our system is logger. This module is responsible for collecting data from users in implicit way, meaning that harvesting of the data does not bother users at all. Four computer mouse events are tracked: mouse movement, single click,

* Supervisor: Daniela Chudá, Institute of Informatics and Software Engineering

Spring 2014 PeWe Workshop, March 21, 2014, pp. 11–12.

mouse wheel movement, scrolling of a page. Each event record is assigned an id of a user, type of the event, time, x and y coordinates of the cursor.

Extractor module converts raw data into meaningful information. Extracted characteristics form the user model that represents user's mouse usage patterns. We calculate tangential velocity, acceleration, angle of the path tangent, angular velocity and curvature of the stroke, a sequence of mouse movement events separated by clicks.

Matcher module provides the identification task itself. The user model constructed in the previous stage becomes the test user model at first. It is compared to all template user models which are stored in the database. The result of matching process is either identity of the test user or the user model becomes template in case no template model matches it and it is stored in the database. The distance between two user models (vectors) is based on t statistics of Welch's t-test.

We conducted a study on 17 users browsing a real running e-shop. An experiment was designed to fully encourage users to perform intended actions by gamification of the user experience. We collected over 90 strokes for each user consisting of at least 4 points. Minimum number of strokes made by one user is 42.

From the view of suitability for identification we examined how distinguishing the values of the characteristics are among the users. To quantify distinctiveness of the characteristic we perform comparison of values for each pair of users using t-test. The distinctiveness could be expressed then as a ratio of number of comparisons indicating difference to all comparisons. The most distinctive characteristic is duration of a single click with distinctiveness rate of 0.77.

Characteristics with higher distinctiveness rate have been selected into the user model. Using the user model containing 17 features we evaluated our identification method and achieved 63.1% success rate of identification.

In our future work we are going to examine changeability of the characteristics over the time and their dependency on hardware used. We are also going to test other classification methods to perform identification task.

Extended version was published in Proc. of the 10th Student Research Conference in Informatics and Information Technologies (IIT.SRC 2014), STU Bratislava, xx-xx.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0971/11.

References

[1] Pusara, M., Brodley, C.E.: User re-authentication via mouse movements. In: *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security - VizSEC/DMSEC '04*, (2004), pp. 1–8.

[2] Gamboa, H., Fred, A.: A behavioral biometric system based on human-computer interaction. In: *Proceedings of SPIE,* (2004).

[3] Zheng, N., Paloski, A., Wang, H.: An efficient user verification system via mouse movements. In: *Proceedings of the 18th ACM conference on Computer and communications security - CCS '11,* (2011), pp. 139-150.