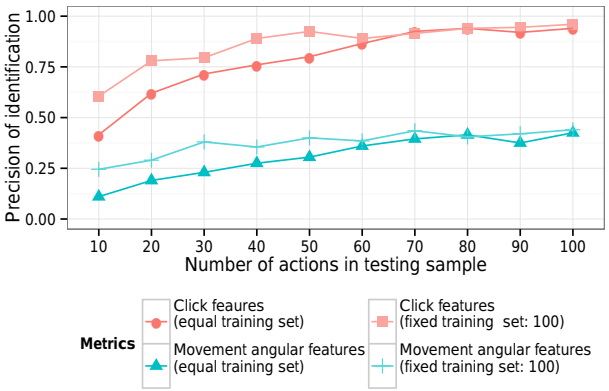

Figure 1: Performance comparison of click and movement features.

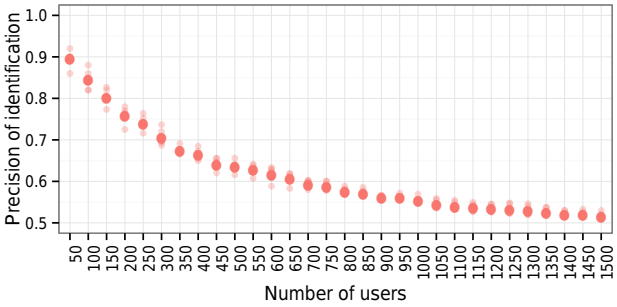

Figure 2: Influence of user pool size on performance.

measures based on mean and deviation are insufficient. We employ statistic of Kolmogorov-Smirnov test to the measure difference of two cumulative distribution functions and the distance between two user models is the sum of the differences for each feature.

4. EXPERIMENTS AND RESULTS

We evaluated the method on two datasets: a controlled experiment with 20 users, and a large scale deployment with 180700 users. To collect data, we implemented a logging (JavaScript) module that records user interface events, including mouse events (mouse down/up, mouse movement) in browser and sends them to server logging backend.

The first dataset was collected in a controlled experiment with 20 users. The users were asked to play a browser memory game, in which cards had to be flipped by mouse clicks. Each of them provided more than 200 click events in ten minutes of game play using the same hardware.

The results in Figure 1 demonstrate that mouse click features are suitable candidates for user recognition. We compared performance of our classification method on the click features and the three features adapted from [5]. The angular features were calculated within each movement stroke (movement bounded with clicks). Initially, we used our method to calculate success rate for both sets of features on exclusive training and testing sets of equal sizes. Additionally, evaluation on a fixed training set (100 samples) was used. The success rate (average of 10 iterations of 2-fold cross validation) using 100 clicks is 96%, while the success rate when using 100 movement strokes is only 44%.

For the second dataset, we collected data from a large Central European tourism portal during a period of 21 days. Total of 2286638 clicks from 180700 unique users (12.7 clicks per user) were collected. Click counts per user followed a long-tailed distribution. For evaluation, we selected users with at least 100 clicks during the collection period, resulting in a subset of 1523 users. For each user, 100 clicks selected randomly were divided into two equal sets for 2-fold cross validations: 50 clicks for training, and 50 clicks for testing (user identification).

The Figure 2 shows how identity recognition success rate (average of 5 iterations of 2-fold cross validation) depends on the size of the user pool. Based on 50 clicks, the recognition accuracy of a user in a pool of 100 users is 85%. On the far end, the accuracy declines slowly. To recognize an unknown user in a pool of 1500 users the success rate is still over 51%. Mean time between two clicks in a session was approximately 23 seconds, requiring a total of 19 minutes to achieve the stated accuracies.

5. CONCLUSIONS

In the paper we present a novel approach to user identity recognition on the web. It shows experiments results benchmarking the classification algorithm based on distance measure adapted from Kolmogorov-Smirnov non-parametric test. The method could be used for accurate recognition of users in small groups (e.g. sharing the same computer) to improve user-oriented services. Additionally, it enables relatively accurate recognition of a user in user pools containing several hundreds of users.

6. ACKNOWLEDGMENTS

This work was partially supported by the grant No. VG 1/0752/14 and grant No. VG 1/0646/15.

7. REFERENCES

[1] D. Chudá and P. Krátky. Usage of computer mouse characteristics for identification in web browsing. In *Proceedings of the 15th International Conference on Computer Systems and Technologies - CompSysTech '14*, pages 218–225, New York, New York, USA, 2014. ACM Press.

[2] H. Gamboa and A. Fred. A behavioral biometric system based on human-computer interaction. In *Proceedings of SPIE 5404, Biometric Technology for Human Identification*, pages 381–392, 2004.

[3] F. Monrose and A. Rubin. Authentication via keystroke dynamics. In *Proceedings of the 4th ACM Conference on Computer and Communications Security*, pages 48–56, 1997.

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

[5] N. Zheng, A. Paloski, and H. Wang. An efficient user verification system via mouse movements. In *Proceedings of the 18th ACM conference on Computer and communications security - CCS '11*, pages 139–150, New York, New York, USA, 2011. ACM Press.