How Human-Mouse Interaction can Accurately Detect Faked Responses About Identity

Merylin Monaro[1], Francesca Ileana Fugazza[2], Luciano Gamberini[1,2], and Giuseppe Sartori[1,2(✉)]

[1] Human Inspired Technology Research Centre, University of Padova, via Luzzati 4, 35122 Padua, Italy
merylin.monaro@phd.unipd.it, {luciano.gamberini,giuseppe.sartori}@unipd.it
[2] Department of General Psychology, University of Padova, via Venezia 4, 35131 Padua, Italy
francescaileana.fugazza@studenti.unipd.it

Abstract. Identity verification is nowadays a very sensible issue. In this paper, we proposed a new tool focused on human-mouse interaction to detect fake responses about identity. Experimental results showed that this technique is able to detect fake responses about identities with an accuracy higher than 95%. In addition to a high sensitivity, the described methodology exceeds the limits of the biometric measures currently available for identity verification and the constraints of the traditional lie detection cognitive paradigms. Thanks to the many advantages offered by this technique, its application looks promising especially in field of national and global security as anti-terrorist measure. This paper represents an advancement in the knowledge of symbiotic systems demonstrating that human-machine interaction may be well integrated into security systems.

Keywords: Identity verification · Lie detection · Mouse tracking

1 Introduction

In the last twenty years, the Global terrorism database (Gtd), the most comprehensive and reliable database on terrorism edited by the University of Maryland, has recorded 70.433 acts of terrorism in the world. Considering the frequency of terrorist attacks from 1994, a rapid growth starting from 2007 to date can be noticed [1]. Due to this alarming increment, a great attention has been paid to the measures currently available to improve the security of nations against terrorist threats.

The report of the National Commission on Terrorist Attacks Upon the United States, also known as 9/11 Commission, suggested the introduction of biometric measures within national borders to prevent the entry of people traveling under false identities [2]. In fact, the use of fake identities is an important means for terrorists because they used false passports to facilitate travel in other countries, such as Europe and US countries. Given this direct link between identity theft and terrorism, the identity verification is a very strong issue directly related to both national and global security [3].

L. Gamberini et al. (Eds.): Symbiotic 2016, LNCS 9961, pp. 115–124, 2017.
DOI: 10.1007/978-3-319-57753-1_10

However, the recognition of terrorists using false identities to move from a country to another is not the only practical context in which identity verification is crucial. The identity verification is a key issue for a large number of application domains, such as the security issue for online authentication (e.g., online banking, ecommerce websites) and the use of fake profiles in social networks.

Biometric measures currently available for identity verification exploit physiological or behavioural characteristics such as fingerprints, hand geometry, and retinas to check identity [4]. More recent approaches developed biometric identification systems based on user-pc interaction characteristics, such as keystroke dynamics and mouse dynamics [5, 6].

Nevertheless in the context of terrorism and in other practical domains, these identity check tools are not useful because many of the suspects are unknown and their biometrical characteristics are not included in databases and, therefore, unidentifiable [7]. For this reason, one actual open challenge is to implement a reliable instrument for identity verification that does not require any prior information about the suspect. In other words, an instrument that recognizes the specific user is not helpful to identify terrorists, thus a tool that detects the deception about identity in a more generic way is necessary.

The deception production is a complex psychological process in which cognition plays an important role [8]. During the generation of a false response, the cognitive system does not simply elaborate a statement, but it carries out several executive tasks: it inhibits the *true* statement and, subsequently, it produces a *false* statement [9]. Moreover, the generation of a lie requires to monitor the reaction of the interlocutor and to adjust the behavior congruently to the lie [10]. All these mental operations cause an increase in cognitive load and, generally, a greater cognitive load produces a bad performance in the task the participant is carrying out, in terms of timing and errors [11]. In particular, participants manifest a lengthening of reaction times (RTs) and an increasing in error rate. This phenomenon has been observed by studying the RTs in double choice tasks: the choice between two alternatives becomes slower in the deceptive response than the truthful one [12].

According to the functioning of our cognitive system, behavior-based lie detection tools have been proposed. The most cited are RT-based Concealed Information Test (RT-CIT) [13] and the autobiographical Implicit Association Test (aIAT) [14] that are two memory detection techniques. Based on RT recording, these instruments can detect between two alternative memories presented to the participant in form of words or sentences which is true and which one is false. These techniques have been used also for identity verification, to reveal which of two identities is the real identity of the examinee [15]. However, both RT-CIT and aIAT require that the true identity information is available, while in the real application only the information provided by suspected is obtainable.

As well as RTs are considered reliable behavioral indices of deception, kinematic analysis of hand movements may provide a clue for recognizing deceits [16]. In fact, recently researchers described as a simple hand movement can be used to study the continuous evolution of the mind processes underlying a behavioural response during a computer task [17].

Applying this evidence to the study of lie, Duran, Dale and McNamara published the results of the first work in which hand movements were used to distinguish deceptive responses to the truthful ones [16]. During the task, participants were instruct to answer *yes* or *no* questions about autobiographical information appearing on a screen using the Nintendo Wii controller. Half of the trials required to response truthfully and the other half required a false response. Results interdicted that deceptive responses could be distinguished from truthful ones based on several dynamic indices, such as the overall response time, the motor onset time, the arm movement trajectory, the velocity and the acceleration of the motion.

Hibbeln and colleagues analysed mouse dynamics in an insurance fraud online context, showing that crafty participants had a different mouse usage pattern in comparison to the honest [18]. The same results have been obtained by Valacich et al. that monitored the mouse activity of fair and guilty people while they were compiling an online survey similar to the Concealed Information Test (CIT) [19].

Based on these pioneering studies, in this paper we propose a new method focused on human-mouse interaction to detect fake responses about identity. The described methodology exceeds the limits both of the traditional RT-lie detection paradigms (e.g., RT-CIT and aIAT) and the biometric measures because any previous information about identity is needed. In fact, the lie detection tool is simply built on the information that an unknown suspect declares. In other words, in this paper we demonstrate how human-machine interaction can improve security, creating a symbiotic system between user and security systems.

2 Method

2.1 Participants

40 participants between students and employees of the Department of General Psychology in Padova University volunteered for this experiment. Participants did not receive any compensation from taking part in the study. All participants agreed on the informed consent. The two experimental group were balanced by gender, age and education (truth-tellers: 10 males and 10 females, mean age = 23.4, mean education = 16.9; liars: 10 males and 10 females, mean age = 25.1, mean education = 16.3).

2.2 Experimental Procedure

The experimental task consisted in 50 double-choice questions about identity in which participants answered clicking with the mouse on the correct alternative response on the computer screen. Half of the participants were instructed to lie about their identities, whereas the 20 control participants answered truthfully. Before the task, the 20 liars learned a fake identity profile from an Italian Identity Card, where a photo of the participant was attached. In order to verify that the information was stored, the fake profile in the ID card was recalled for two times, interspersed with a mathematical distracting task. Truth tellers performed a mathematical task and revised their real autobiographical data only once before starting the experiment.

The experiment was implemented and run on a laptop (15.6″) using *MouseTracker* software [20]. Six practice questions preceded the experimental task. Questions appeared centrally in the upper part of the computer screen. The response labels were located one on the right and one on the left upper bound of the screen. Response labels appeared at the same time of the question.

Table 1. The table reported some examples of presented questions to the participants and the possible answers.

Type of question	Example of correct response	Example of incorrect response
Control questions		
Are you female?	Yes	No
Are you male?	No	Yes
Do you have any tattoos?	No	Yes
Do you have pierced ears?	Yes	No
What is your shoe size?	36	42
What is your eye color?	Brown	Blue
How tall are you?	160 cm	190 cm
What is your skin color?	White	Black
Expected questions		
Were you born in April?	Yes	No
Were you born in October?	No	Yes
Do you live in Padova?	Yes	No
Do you live in Napoli?	No	Yes
What is your last name?	Moretti	Greco
What is your year of birth?	1987	1984
What is your city of birth?	Verona	Milano
What is your name?	Sara	Anna
Unexpected questions		
Are there any double letters in your last name?	Yes	No
Do you live in the same region where you were born?	No	Yes
Is your residence city near Abano Terme?	Yes	No
Is your residence city near Saturnia Terme?	No	Yes
How old are you?	28	25
Which is your zodiac sign?	Aries	Capricorn
What is your zip code?	35142	36125
What is the chief town of your born region?	Venezia	Firenze

The half of the questions requested a *yes* or *no* response, while the other requested a response to different labels (e.g., to the question "*Which is your gender?*" possible

response labels might be "*male*" "*female*"). Within the entire task, the correct responses, that are the answers congruent with the suspect declarations, were presented for the 50% of trials on right position and for the other 50% on the left. Some examples of the 50 questions included in the experimental task are reported in Table 1.

During the experiment, three different kinds of questions were randomly presented to participants. *Expected questions* were information that has been learned by liars from the fake ID card and explicitly trained during the learning phase (e.g., "*Were you born in 1987?*"), whereas *unexpected questions* derived from this information but were not explicitly rehearsed before the experiment (e.g., "*Are you 29 years old?*"). Finally, *control questions* required a true response both for liars and for truth-tellers because they concerned physical information, which is not possible to hide (e.g., "*Are you female?*"). As reported in literature, the presence of *unexpected questions* has the effect to increase the cognitive load in liars [21]. Whereas for truth-tellers the unexpected information is quickly and easily available even if they are not prepared to those specific questions, liars have to fabricate a new response congruently with the other ones. Because this mental operation requires a greater cognitive effort, liars show in *unexpected questions* a bad performance compared with truth-tellers.

2.3 Collected Measures

During each response, the *MouseTracker* software recorded the following kinematic features:

- *X,Y coordinates over the time* (X_n, Y_n): position of the mouse along the axis over the time. Because each trajectory has a different length, in order to permit averaging and comparison across multiple trials, the *MouseTracker* normalizes each motor response in 101 time frames [20].
- *Velocity over the time* (vX_n, vY_n): velocity of the mouse along the axis over the time.
- *Acceleration over the time* (aX_n, aY_n): acceleration of the mouse along the axis over the time.
- *Initiation time* (IT): time between the appearance of the question and the beginning of the response.
- *Reaction time* (RT): time between the appearance of the question and the end of the response.
- *Maximum deviation* (MD): largest perpendicular distance between the actual trajectory and the ideal trajectory.
- *Area under the curve* (AUC): geometric area between the actual trajectory and the ideal trajectory.
- *Maximum deviation time* (MD-time): time to reach the point of maximum deviation.
- *x-flip*: number of direction reversals along the *x*-axis.
- *y-flip*: number of direction reversals along the *y*-axis.
- *Number of errors*: number of incorrect responses.

For each feature we calculated the mean value within participants for all trials. Finally, we used these values to perform statistical analysis and to build a machine learning classification model.

3 Analysis and Results

3.1 Graphical Observations and Statistics

We graphically compared the performance of the two experimental groups (liars vs truth-tellers), separately for *control*, *expected* and *unexpected questions*. Figure 1 reports the average trajectories for liars and truth-tellers, respectively for *control*, *expected* and *unexpected* questions. Furthermore, the figures below represent the average position of the mouse on *x* and *y*-axis over the time. As it can be noticed, the trajectories of the two experimental groups visually differ especially for the *unexpected questions*, whereas for the *control* and the *expected questions* this difference is not so evident. Considering *unexpected questions*, the truth-teller response shows a more direct trajectory, connecting the starting point with the end-response point. By contrast, liars spend more time moving on *y*-axis in the initial phase of the response and deviate to the selected response with a certain delay compared to truth-tellers.

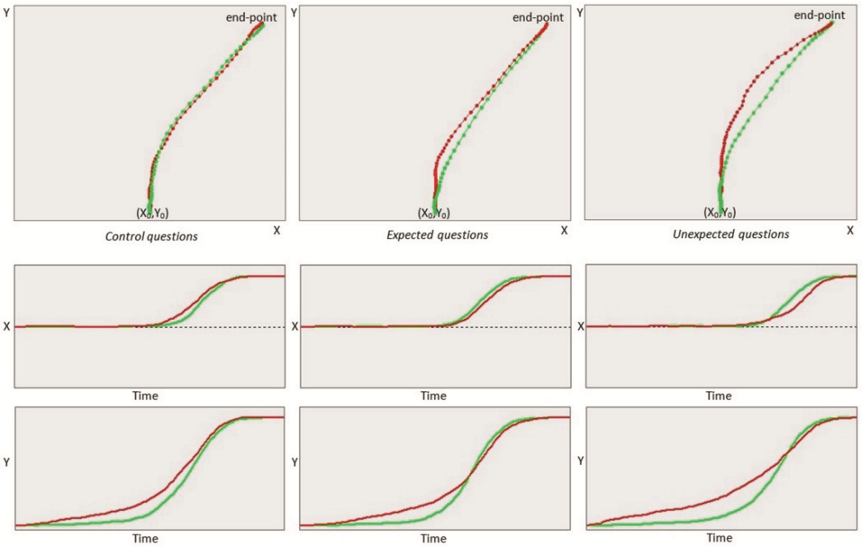

Fig. 1. The panels displayed in the first row report, separately for *control*, *expected* and *unexpected questions*, the average trajectories for liars (red line) and truth-tellers (green line). The panels in second and third row show the average position of the mouse on *x* (second row) and *y*-axis (third row) over the time for liars (red line) and for truth-tellers (green line), respectively for *control* (left panel), *expected* (central panel) and *unexpected questions* (right panel). In other words, these panels represent how the mouse moves along the *x* and *y*-axis during the 101 response time frames. (Color figure online)

In order to confirm whether the difference between liars and truth-tellers trajectories in *unexpected questions* is statistically significant, we run an independent t-test on the collected measures (see Subsect. 2.3). Results showed that liars' responses significantly differ from truth-tellers' ones in AUC ($t = 3.13$, $p < 0.0042$), RT ($t = 3.61$, $p < 0.0042$),

and mean velocity along x-axis ($t = -7.62$, $p < 0.0042$). Finally, liars make a higher number of errors compared to truth-tellers ($t = 9.70$, $p < 0.0042$) (to avoid the multiple testing problem the correction of Bonferroni has been apply and the p-value has been set to 0.0042).

Finally, we tested the difference between liars and truth-tellers also for *expected* and *control questions*, confirming that none of the measures considered (see Subsect. 2.3) reach the statistical significance in the independent t-test.

3.2 Machine Learning Models

According to the results, obtained by graphical and statistics observations, we used only *unexpected questions* data to train different machine learning classifiers. The goal is to create a model that is able to predict whether the participant is a liar or a truth-teller, based on the mouse response features. To optimize the accuracy of our model, we perform a feature selection, according to the attribute selection function that is implemented in WEKA software [22]. In particular, we run a ranker analysis [23]: this function uses an attribute/subset evaluator to rank all attributes inserted in the model as predictors. The ranked list of the 12 features considered (see Subsect. 2.3) is the following: errors = 0.83, mean velocity along x-axis = 0.62, AUC = 0.24, RT = 0.23, MD = 0.2, all the other features = 0.00. It can be noticed that the features that show a greater weight for the model, according to the ranker analysis, are the same that reached a significant t-value in the independent t-test (see above). For this reason, we decided to select these four features to implement the classification models.

The classification procedure has been performed using WEKA software [22].

Classification models have been built using a 10-fold cross-validation procedure as implemented in WEKA. Table 2 reports the percentage accuracy values of the different classifiers. It can be noticed that the classification accuracy remains stable across the different classifiers, ranging from 90% to 95%.

Table 2. The table reports the accuracy values in the 10-fold cross-validation for four different ML classifiers: Simple Logistic [24], Support Vector Machine (SVM) [25], Random forest [26] and Naive Bayes [27]. The classification accuracy is reported considering all *unexpected questions*, questions requiring a *yes* or *no* response and questions requiring a response to *different* labels.

Classifier	Accuracy in 10-fold cross validation for all *unexpected questions*	Accuracy in 10-fold cross validation for *unexpected questions* requiring a *yes* or *no* response	Accuracy in 10-fold cross validation for *unexpected questions* requiring a response to *different* labels
Simple Logistic	90%	80%	100%
SVM	95%	77.5%	97.5%
Random Forest	90%	75%	100%
Naive Bayes	95%	67.5%	97.5%

Finally, we separately repeated the classification procedure for the *unexpected questions* that required a *yes* or *no* response and for questions that requested a response to different categories labels (e.g., "*male*" "*female*"). The percentages of accuracy are reported in Table 2. Results show that, considering only the questions that require a response to different labels, the classification accuracy improves from 2.5% to 10%. In other words, we reach the best classification performance in distinguishing liars from truth-tellers on *unexpected questions* that requested a response to different categories labels.

4 Discussion and Conclusions

In this work, we described a new tool to detect liars about identity. The technique exploits the user-mouse interaction when the suspect is engaged in a computerize task requiring identity information. We tested the method through an experiment involving 40 participants. Half of participants was instructed to declare a fake identity according to a false ID card previously learned. Then, questions about identity information (e.g., name, surname, date of birth, etc.) were presented. Participants clicked with the mouse on the correct response between the two alternatives, according to the identity information that they declared. *Unexpected questions* were introduced to increase the cognitive load in liars. Moreover, we introduced a variability in response labels. In other words, participants did not answer only to fixed *yes* or *no* questions but to different categories questions (e.g., to the question "*How old are you?*" possible response labels might be "*25*" "*28*").

The kinematic features of the motor response were collected and used to train different machine learning classifiers.

Responses to unexpected questions are those in which, according both with graphical and statistical observations, liars and truth-tellers show the main difference.

The accuracy, obtained by the classification models in correctly predicting the veracity of the declared identity, is very high, around 95%. Nevertheless, we point out that to confirm the stability of our model and in order to ensure the reproducibility of the data, it will be needed to extend the number of observations included in the training set and to collect a further sample of naïve participants to test the model with an out-of-sample procedure [28].

Results also showed that, considering only *unexpected questions* that required a response to different categories labels, the accuracy improves to 97.5–100%.

Our hypothesis is that the continuous change of the response label categories results in a further increase in cognitive load for liars. In fact, it is possible that using only *yes* or *no* fixed labels, after some trials the label processing becomes partially automated and does not require any mental effort. Conversely, in a task where label categories change away, the true label is often very familiar for truth-tellers. However, the true and the false label are unfamiliar for liars, especially in the case of unexpected questions. For this reason, the liars' response requires more cognitive effort to process labels and implement the correct response. This effort causes a deterioration of the liars' performance and the discrimination between the two experimental groups becomes more accurate. Furthermore, it is possible that the classification accuracy in questions requiring a

yes or *no* response is lower because liars answer falsely to questions requiring a *yes* response, but they are truthful in answering questions required a *no* response.

In conclusion, this paper represents an advancement in the knowledge of symbiotic interaction demonstrating that human-computer interplay may improve security systems, creating a symbiosis between user and security. This methodology seems to be promising in detecting fake responses about identity for several reasons. In addition to the high accuracy, one of the most innovative advantages of this tool is that it does not requires any knowledge about the real identity of the suspect. Secondly, the classification algorithm exploits a large number of kinematic indices to identify liars, so it is difficult to control via efficient countermeasures all these parameters. Finally, it is cheap both in terms of money and in terms of time for testing. This last feature makes it suitable for large-scale applications, as the control of the international migration flow.

References

1. National Consortium for the Study of Terrorism and Responses to Terrorism (START). Global Terrorism Database [Data file] (2015). http://www.start.umd.edu/gtd
2. National Commission on Terrorist Attacks Upon the United States. http://govinfo.library.unt.edu/911/report/index.htm
3. The University of Texas at Austin (2015). http://news.utexas.edu/2015/12/07/the-direct-link-between-identity-theft-and-terrorism
4. Ashbourn, J.: Biometrics: Advanced Identity Verification: The Complete Guide. Springer, Heidelberg (2000)
5. Shen Teh, P., Beng Jin Teoh A., Yue, S.: A survey of keystroke dynamics biometrics. Sci. World J. **24** (2013)
6. Feher, C., Elovici, Y., Moskovitch, R., Rokach, L., Schclar, A.: User identity verification via mouse dynamics. Inf. Sci. **201**, 19–36 (2012)
7. Sabena, F., Dehghantanha, A., Seddon, A.P.: A review of vulnerabilities in identity management using biometrics. In: Second International Conference on Future Networks, ICFN 2010, Sanya, Hainan, pp. 42–49 (2010)
8. Abe, N.: How the brain shapes deception: an integrated review of the literature. Neuroscientist **17**, 560–574 (2011)
9. Debey, E., De Houwer, J., Verschuere, B.: Lying relies on the truth. Cognition **132**, 324–334 (2014)
10. Gombos, V.A.: The cognition of deception: the role of executive processes in producing lies. Genet. Soc. Gen. Psychol. Monograhs **132**, 197–214 (2006)
11. Blandon-Gitlin, I., Fenn, E., Masip, J., Yoo, A.H.: Cognitive-load approaches to detect deception: searching for cognitive mechanisms. Trends Cogn. Sci. **18**(9), 441–444 (2014)
12. Sheridan, M.R., Flowers, K.A.: Reaction times and deception - the lying constant. Int. J. Psychol. Stud. **2**(2), 41–51 (2010)
13. Kleinberg, B., Verschuere, B.: Memory detection 2.0: the first web-based memory detection test. PLoS ONE **10**(4), e0118715 (2015)
14. Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S.D., Castiello, U.: How to accurately detect autobiographical events. Psychol. Sci. **19**(8), 772–780 (2008)
15. Verschuere, B., Kleinberg, B.: ID-check: online concealed information test reveals true identity. J. Forensic Sci. **61**, S237 (2015)

16. Duran, N., Dale, R., McNamara, D.S.: The action dynamics of overcoming the truth. Psychon. Bull. Rev. **17**(4), 486–491 (2010)
17. Freeman, J.B., Dale, R., Farmer, T.A.: Hand in motion reveals mind in motion. Front. Psychol. **2**, 59 (2011)
18. Hibbeln, M., Jenkins, J., Schneider, C., Valacich, J., Weinmann, M.: Investigating the effect of fraud on mouse usage in human-computer interactions. In: International Conference on Information Systems, ICIS 2014 (2014)
19. Valacich, J.S., Jenkins, J.L., Nunamaker, Jr., J.F., Hariri, S., Howie, J.: Identifying insider threats through monitoring mouse movements in concealed information tests. In: Hawaii International Conference on System Sciences. Deception Detection Symposium (2013)
20. Freeman, J.B., Ambady, N.: MouseTracker: software for studying real-time mental processing using a computer mouse-tracking method. Behav. Res. Meth. **42**, 226–241 (2010)
21. Lancaster, G.L.J., Vrij, A., Hope, L., Waller, B.: Sorting the liars from the truth tellers: the benefits of asking unanticipated questions on lie detection. Appl. Cogn. Psychol. **27**, 107–114 (2013)
22. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. ACM SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
23. Hall, M., Holmes, G.: Benchmarking attribute selection techniques for discrete class data mining. IEEE Trans. Knowl. Data Eng. **15**(6), 1437–1447 (2003)
24. Landwehr, N., Hall, M., Frank, E.: Logistic model trees. Mach. Learn. **95**(1–2), 161–205 (2005)
25. Keerthi, S.S., Shevade, S.K., Bhattacharyya, C., Murthy, K.R.K.: Improvements to Platt's SMO algorithm for SVM classifier design. Neural Comput. **13**(3), 637–649 (2001)
26. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
27. John, G.H., Langley, P.: Estimating continuous distributions in Bayesian classifiers. In: Eleventh Conference on Uncertainty in Artificial Intelligence, San Mateo, pp. 338–345 (1995)
28. Dwork, C., et al.: The reusable holdout: preserving validity in adaptive data analysis. Science **349**, 636–638 (2015)