

Social Engineering Detection Using Behavioral Data and Form-Usage Patterns

Yuvaraja P¹ and Sathyanarayanan R²

¹Department of Artificial Intelligence and Data Science, St Joseph's Institute of Technology
,Chennai, Tamil Nadu

²Department of Artificial Intelligence and Data Science, St Joseph's Institute of Technology
,Chennai, Tamil Nadu

Abstract. Social engineering attacks capitalize on the human mental vulnerability and not technical vulnerability and thus it is not easily spotted by traditional cybersecurity controls. Traditional security mechanisms like authentication system, phishing system, and intrusion system detection systems are largely centered on infrastructure based threats and in most cases fail when genuine users are compromised on authenticated systems. This paper will suggest a real-time behavioural monitoring model to be used in identifying possible manipulation of social engineering during web-form communications. The solution will combine behavioral biometrics and contextual form-usage analysis to detect any interaction anomalies that can suggest cognitive manipulation. Lightweight client-side monitoring captures behavioral information such as the timing of keystroke, the behavior of mouse movements, pauses, and patterns of corrections, and contextual information such as dwell time, field-entry sequence deviation, and frequency of re-edits are derived out of form interaction behavior. The features are converted to session-level vectors and tested based on machine learning models. Experimental assessment of 50 subjects demonstrates that Multi-Layer Perceptron model finds 94.1% percentage with an AUC of 0.96 with less than 1-second inference delay.

1 Introduction

The high growth of online platforms and cloud services, has brought a great revolution in contemporary computing systems. The web based infrastructures are becoming essential in the manner in which online banking, e-commerce services, healthcare systems and enterprise applications are used to carry out their daily operations. Despite the accessibility and efficiency offered by these technologies, these technologies also increase the attack surface that a cyber threat can attack. Over the last years, attackers have started paying less attention to utilizing software vulnerabilities and more to the manipulation of human users, which has provided more and more social engineering attacks. These are psychological attacks that use urgency, the pressure of authority, fear and trust exploitation as their approaches to influence the users to divulge confidential information or take unintended actions. Such attacks, as they are aimed at human behavior, tend to evade the traditional cybersecurity measures.

Conventional security systems including password, multi-factor, encryption, and phishing detection systems are basically set up to detect unfriendly software executions or unauthorized access attempts. Nevertheless, such mechanisms fail to work in many instances where a

legitimate user is swindled over in an authenticated session. Under these circumstances, the system will monitor legitimate credentials and standard access practices and the user will obliviously undertake activities that will jeopardize security. This leads to a growing requirement of mechanisms of detection that are aimed at detection of not only system-level anomalies but also deviations in behavior of users when interacting.

Behavioral biometrics has become a prospective and effective method of studying human interaction with the computing systems. The studies of human-computer interaction have found that behavioral cues (typing rhythm, mouse movement pattern, hesitation period, and cursor dynamics) have characteristic features that can indicate changes in the cognitive state of the user. Once the users are exposed to stress, confusion or external pressure which in most cases can be linked with social engineering manipulation, the interaction pattern is likely to be affected unlike normal behavior. Such deviations give useful tips, which are applicable in identifying any manipulation attempts made by the web-based activities.

This paper was inspired by these findings and aims to come up with a real-time social engineering attack detection behavioral analytics system when interacting

with web forms. The given solution is based on behavioral biometric cues with the contextual patterns of form-usage so as to detect interactions-level anomalies that might reflect cognitive manipulation. The system will attempt to warn of a potential leakage of sensitive information by observing the actions of the users at the time of the input session and by using machine learning models to identify patterns of interactions, which will then aid in triggering an immediate response to the signs of information leakage before the form is submitted.

2 Literature Survey

Social engineering emerged as one of the greatest dangers of the current-day cybersecurity since it is based on psychological vulnerability of humans, but not the technical vulnerability of the computing systems. The conventional defense protection (firewalls, intrusion detection systems, authentication protocols, etc.) is mostly concerned with infrastructure and network layer protection. Inasmuch as these mechanisms are useful in combating malware attacks and unauthorized access to the system, they are not as effective in detecting a scenario where a normal user is manipulated to willingly release a sensitive information. Research on phishing and social manipulation attacks has shown that even the technically secure systems can be undermined when an attacker uses his or her human trust, urgency, or authority to shape the behavior of the users.

Other studies conducted in the past in phishing detection have centered mostly on detecting malicious websites, emails or URLs through machine learning and pattern recognition methods. Webpage content analysis, visual similarity detection, and domain classification have proved to be promising in the detection of a fraudulent platform. Nonetheless, these methods are not applicable in a situation where attackers are working in real systems or manipulate the users to execute activities in the authenticated sessions. When this happens, the traditional detection systems may have nothing malicious or suspicious in the network to detect.

Behavioral biometrics has become another approach in improving cybersecurity through the interaction of users with the computing devices. Continuous authentication and fraud detection have extensively used techniques like keystroke dynamics, analysis of mouse movements and typing rhythm modeling. Studies have revealed that patterns of neuromuscular interaction such as inter-key latency, cursor acceleration and typing speed, were relatively constant among individual users. These attributes enable the behavioral biometrics to be an implicit authentication regarding identity verification when using a system. However, the available behavioral

biometric systems are mainly used to identify people, instead of identifying cognitive manipulation during communication.

Human-computer interaction research also gives additional information on how cognitive stress and external pressure can affect the user behavior. Empirical research shows that the people exposed to stress or uncertainty tend to show experimentally significant variation in patterns

System Architecture of Social Engineering Detection Framework

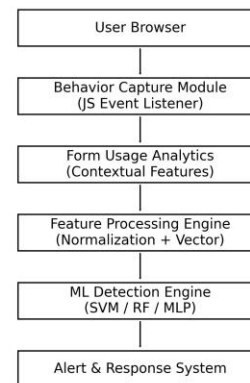


Figure 1: System architecture of the proposed social engineering detection framework.

of interaction, including a longer hesitation period, nonlinearity in navigation routes, a longer stay time on input fields, and repetitive correction behavior. Such behavioral anomalies are indications that interaction-level analytics can be of good use in identifying scenarios where users can be manipulated or coerced. Irrespective of the results, very few studies have delved into the domain of integrating behavioral biometrics and contextual interaction analytics in detecting social engineering attacks in real time. The study fills this gap by suggesting a framework that integrates the response of signals of behavioral interaction with the contextual analysis of form-usage in order to identify cognitive manipulation in web-based user interactions.

3 System Architecture

The proposed system is planned to be a real-time client server system which tracks the behaviour of user interaction during the use of web-forms and determine the possible signs of manipulation in the social engineering. The architecture incorporates behavioral monitoring and contextual analytics coupled with the machine learning-based detection to examine the patterns of interaction

during form completion. The system works by gathering the metadata of the interaction of the user on the browser, converting the events gathered into structured feature representations and assessing the resultant behavioral patterns with trained classification models. Through this architecture, the detection process is continuous throughout the interaction session without disruption to the normal user working process.

The client side of the system is in charge of gathering behavioral cues as the user interacts with a web interface. The web application will have lightweight JavaScript event listeners that track interaction events including key presses, movement paths of the mouse, changes in the velocity of the cursor, changing focus of fields, and pastes. Instead of capturing the literal textual content typed by the user, the system instead captures solely behavioral metadatas such as timing spans, frequency of events and pathways to ensure the privacy of the user. These uncoded events of interaction are sent safely to the backend processing environment where they are consolidated into session level data streams.

An event log processing (ELP) module running on the server side operates on the gathered event logs converting the interaction events into structured numerical feature vectors. They use noise filtering, session segmentation, normalization and feature scaling to ensure that there is consistency across interaction sessions. The derived feature vectors constitute behavioral and contextual attributes of the interaction process and they are fed into the machine learning detection engine. Each session is tested by the detection engine with trained classification models according to which the probability of cognitive manipulation with regard to observable behavioral deviations is estimated. Once the risk score has passed a pre-determined threshold, a system response mechanism can be activated that can lead to warning notifications, further authentication, or temporarily block form submission. This architecture allows to detect suspicious interaction behavior in real time with minimal computational load and the privacy of user input data is not compromised.

4 Methodology

The suggested methodology is aimed at determining behavioral deviation which can arise during the interaction of users with web forms in the conditions of normal functioning and in the conditions of simulated effects of social engineering. The methodology involves the combination of behavioral biometrics and contextual interaction analytics to ensure that the user action on the motor level and the level of interaction are both captured. The general workflow is initiated by controlled data

gathering, feature extraction, feature preprocessing, and machine learning based classification. The system can form behavioral representations, based on the information gathered throughout the interaction session, and not on discrete events, and that can reflect the cognitive state of the user when filling in the form.

The experimental data were also obtained by holding experimental interaction sessions with fifty people. All participants were given several web-form tasks that would mimic real activities on the internet like the registration of accounts and submitting of information. There were two categories of sessions that were captured: ordinary interaction sessions and the manipulated sessions which were created to replicate the social engineering influence. During manipulated sessions, the participants were introduced to scripted prompts that depicted typical social engineering situations, e.g., an emergency need, instructions by authority, or false guidance. These prompts were

| Feature | Description | Measurement |
|--------------------------|--|-------------------------|
| Typing Speed | Rate of keyboard input during form interaction | Characters/sec |
| Inter-Key Latency | Time between consecutive key presses | Milliseconds |
| Backspace Frequency | Number of correction actions during typing | Count/session |
| Mouse Velocity | Average cursor movement speed | Pixels/sec |
| Mouse Acceleration | Change in cursor velocity during movement | Pixels/sec ² |
| Hesitation Time | Delay before entering information in fields | Seconds |
| Field Dwell Time | Time spent on each input field | Seconds |
| Field Sequence Deviation | Deviation from expected form entry order | Pattern deviation |
| Re-Edit Frequency | Number of times fields are modified | Count/session |

Methodology Workflow for Social Engineering Detection

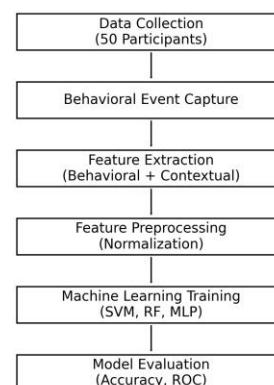


Figure 2: Methodology workflow for behavioral social engineering detection.

provided as messages or verbal instructions to simulate real conditions of manipulation. The final dataset included equal samples of both control and tampered interaction sessions so that it could allow the support of sound models training and testing.

In every session, a behavioral capture module on the client-side captured fine-grained interaction events created by the user. Timing of key presses, duration of key hold, cursor movement patterns, mouse speed changes, frequency of clicks, and changing input field were all monitored by the system. Since the system is concerned with the patterns of behavior as opposed to the analysis of the contents, the text typed by the user is not stored. Rather it documents the temporal and statistical features of user interactions only. These raw events were converted into behavioral indicators which included typing speed, variation in the inter-key latency, frequency of the back space, hesitation duration, the variance of mouse acceleration, and irregularities in the movements of the cursor. These measures of behavioural consistency are indicators of motor-level consistency of user interactions and can show some evidence of cognitive stress or indecision.

Besides motor-level behavioral indicators, the contextual interaction features were obtained based on the structure and usage patterns of the web form. Such contextual cues are the time spent on fields of input, inconsistency with anticipated order of field-entry, recidivism of fields that have already been filled in, and copy-pasting in sensitive fields. Studies on human-computer interaction indicate that cognitive pressure or manipulation tends to cause inconsistency in the sequence of interaction, as well as heightened reluctance in making decisions. Through intersectional interaction signatures and behavioral biometrics, the system can gain a more detailed view of the activity of the users.

The behavioral and contextual features were obtained and concatenated into session-level feature vectors, which represented every interaction session. Before the training of the models, preprocessing methods such as noise filter, normalization and feature scaling were used to minimize the variability of the sessions. The trained dataset was split into an eighty and twenty split to form the training and testing dataset. Several supervised machine learning algorithms were tested (Random Forest, Support Vector Machine, Logistic Regression and Multi-Layer Perceptron). The cross-validation was used to evaluate model performance in order to minimize overfitting and guarantee the ability to generalize. The trained models were then applied to classify interaction sessions either as normal or possibly manipulated allowing the system to make an estimate of the probability of being influenced by

social engineering when engaged in a real time web interaction.

5 Result And Discussion

The performance of the social engineering detection framework that was proposed was measured in terms of standard classification performance measurements such as accuracy, precision, recall, and F1-score. Those metrics will present the overall evaluation of the model to identify the manipulated interaction sessions correctly and the low rate of false detection. The experimental data were behavioral interaction sessions recorded with fifty participants and the models were optimized and tested on an eighty to twenty training and testing split with five fold cross-validation in order to enhance the reliability.

The Multi-Layer Perceptron model was the most successful among the assessed models. The model has the overall accuracy 94.1 and a precision of 93.3, recall of 95.2, and F1-score 94.2. The recall value is very high, so it means that the model was capable of identifying most of the manipulated interaction sessions; this is especially crucial in security-related applications where the inability to detect

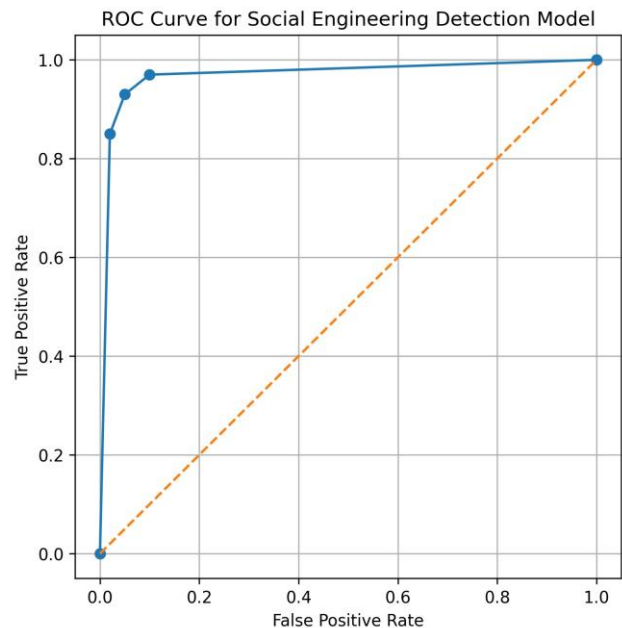


Figure 3: ROC curve illustrating classification performance of the detection model.

an attack can result in sensitive information leakage. The classification with the Random Forest also delivered good results with an overall accuracy of 92.4 and this suggests that the ensemble learning methods are very useful in modeling convoluted relationships between behavioral

and contextual interaction attributes. Comparatively, Support Vector Machine has been seen to have a high level of accuracy of 89.7% and the Logistic Regression got a lower mark because its decision boundary is linear and as such, fails to capture all nonlinear relationship in the behavioral interaction data.

Analyzing the behavior of manipulated sessions revealed that there are some interaction deviations that showed consistency. Simulated social engineering situations led to a longer hesitation time until sensitive information was typed in by the users, more backspace keys were typed, and the users took up uneven routes in maneuvering the form fields. Besides, manipulated sessions were characterized by more discrepancy in dwell time on certain input fields and a higher rate of re-editing of information entered earlier. The changes in behavior are consistent with the results of the human-computer interaction studies that indicate that cognitive stress and external pressure can be a significant factor that determines user interaction behavior. Integrating behavioral biometrics with contextual form-usage analytics, the system could more reliably detect these deviations than the methods that used only one type of features.

The Receiver Operating Characteristic analysis was also used to further test the discriminative capability of the detection model. The ROC curve shows the trade-off of the true positive rate versus the false positive rate in the various levels of classification. The area under the curve factor

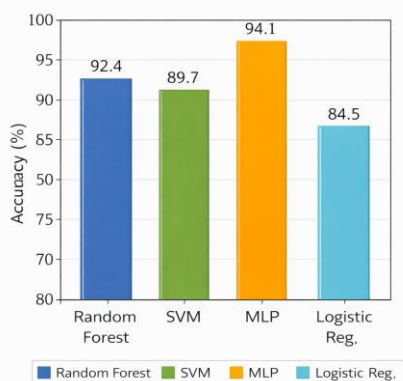


Figure 4: Machine learning model accuracy comparison

of the proposed detection model was 0.96 which showed good classification capability and reliability in terms of distinguishing normal interactivity and manipulated sessions. Besides the performance in terms of classifier classification, the system latency was measured to assess the possibility of the real-time deployment.

Inference times per session classification were under a second on average, proving that the framework can be used in a real-time setting, and will not introduce delays that are noticeable by users. The findings affirm the fact that behavioral interaction monitoring offers a viable and efficient platform of identifying possible social engineering attacks in the process of web-based actions.

6 Conclusion

The problem of social engineering attack is still a major menace in contemporary cyber security, as it takes advantage of cognitive vulnerability of humans as opposed to technical vulnerability of computer systems. The conventional security measures, which include authentication, encryption, and phishing detection systems, are mainly aimed at revealing the infrastructure level threats and malware programs. Nevertheless, such strategies fail to work when authorized users are compromised to carry out malicious activities in the course of authenticated sessions. Consequently, there is the increased demand of detection mechanisms that can be used to detect the behavioral abnormalities that may arise during user engagement with web systems.

This paper introduced a live behavioral analytics model that aims at identifying possible social engineering manipulation in a web-form interaction. The suggested solution includes the combination of behavioral-biometrics and form-usage contextual analytics to ensure the identification of both motor-level feedback of interactions and higher-level behavioral patterns. The system generates the session-level representations that can be seen as the cognitive state of users when completing the form by analyzing the rhythm of typing, the dynamics of the movement of the mouse, the hesitation period, the dwell time, and navigation behavior. These patterns of interactions are assessed with the help of supervised machine learning models to consider the sessions as normal or possibly manipulated.

The proposed framework was tested based on as much as fifty participants with the help of the interaction data, which proved that the suggested framework is effective in terms of identifying behavioral deviations linked to the impact of social engineering. The Multi-Layer Perceptron was the most successful model which had 94.1% accuracy in its classification that had high recall and F1-score scores. The robustness of the detection technique was additionally proven with the help of Receiver Operating Characteristic analysis, which also provided the Area Under Curve of 0.96. Moreover, the system still had a sub-second inference latency meaning that the suggested framework can be deployed in real-time in web-based applications.

Even though the results are showing good performance, there are some limitations. The experimental data were gathered under a controlled setting and had a small number of subjects so it might not be a complete representation of the diversity in reality with regard to user behavior. The future studies will include widening the sample base to include larger and more varied groups and assessing the system robustness to adversarial behavioral mimicry and extending the model to cover the mobile and voice interaction setting. On the whole, behavioral monitoring on the interaction level offers a scalable and non-invasive method of social engineering attack detection and increasing the security of human-centered digital systems.

References

- [1] A. Jain and B. Gupta, "Phishing detection: Analysis of visual similarity-based approaches," *Journal of Information Security and Applications*, vol. 36, pp. 68–81, 2017.
- [2] Y. Alsharnouby, F. Alaca, and S. Chiasson, "Why phishing still works: User strategies for combating phishing attacks," *International Journal of Human-Computer Studies*, vol. 82, pp. 69–82, 2015.
- [3] D. Bojinov, E. Bursztein, D. Boneh, and P. Bursztein, "Using behavioral biometrics for continuous authentication in real-world settings," in *Proc. IEEE Security and Privacy Workshops*, pp. 187–193, 2014.
- [4] J. Acién, A. Morales, R. Vera-Rodriguez, J. Fierrez, and R. Tolosana, "BeCAPTCHA: Behavioral bot detection via mouse dynamics," *arXiv preprint arXiv:2005.13655*, 2020.
- [5] S. Kumar, A. Arora, and R. Sharma, "Hybrid anomaly detection using mouse and keystroke dynamics for phishing prevention," *International Journal of Cybersecurity and Digital Forensics*, vol. 11, no. 2, pp. 77–87, 2022.
- [6] L. F. Cranor, "Security warning fatigue: A case study," in *Proc. 26th Annual Computer Security Applications Conference (ACSAC)*, pp. 121–130, 2016.
- [7] S. Chiasson, P. C. van Oorschot, and R. Biddle, "A usability study and critique of two phishing user studies," in *Proc. 15th USENIX Security Symposium*, pp. 1–14, 2006.
- [8] C. Whittaker, B. Ryner, and M. Nazif, "Large-scale automatic classification of phishing pages," in *Proc. Network and Distributed System Security Symposium (NDSS)*, 2010.
- [9] K. Snow and F. Calabrese, "Behavioral fingerprinting for intrusion detection," in *Proc. IEEE Conference on Communications and Network Security (CNS)*, pp. 493–498, 2014.
- [10] A. Moghimi, A. B. Nassif, and R. M. A. Abdullah, "Machine learning for detecting social engineering attacks: A review," *IEEE Access*, vol. 10, pp. 12098–12115, 2022.
- [11] A. Aleroud and L. Zhou, "Phishing environments, techniques, and countermeasures: A survey," *Computers & Security*, vol. 68, pp. 160–196, 2017.
- [12] M. Jakobsson and S. Myers, *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*, Wiley-Interscience, Hoboken, NJ, USA, 2006.
- [13] J. Bonneau, C. Herley, P. C. van Oorschot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes," in *Proc. IEEE Symposium on Security and Privacy*, pp. 553–567, 2012.
- [14] F. Monroe and A. Rubin, "Keystroke dynamics as a biometric for authentication," *Future Generation Computer Systems*, vol. 16, no. 4, pp. 351–359, 2000.
- [15] R. Biddle, S. Chiasson, and P. C. van Oorschot, "Graphical passwords: Learning from the first twelve years," *ACM Computing Surveys*, vol. 44, no. 4, pp. 1–41, 2012.
- [16] T. Jagatic, N. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Communications of the ACM*, vol. 50, no. 10, pp. 94–100, 2007.
- [17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- [18] P. Legg, O. Buckley, M. Goldsmith, and S. Creese, "Automated insider threat detection system using user and role-based profile assessment," *IEEE Systems Journal*, vol. 11, no. 2, pp. 503–512, 2017.