Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2023.036020
Article

Grey Wolf-Based Method for an Implicit Authentication of Smartphone Users

Abdulwahab Ali Almazroi and Mohamed Meselhy Eltoukhy*

University of Jeddah, College of Computing and Information Technology at Khulais, Department of Information Technology, Jeddah, Saudi Arabia
*Corresponding Author: Mohamed Meselhy Eltoukhy. Email: mmeltoukhy@uj.edu.sa
Received: 14 September 2022; Accepted: 29 January 2023

Abstract: Smartphones have now become an integral part of our everyday lives. User authentication on smartphones is often accomplished by mechanisms (like face unlock, pattern, or pin password) that authenticate the user's identity. These technologies are simple, inexpensive, and fast for repeated logins. However, these technologies are still subject to assaults like smudge assaults and shoulder surfing. Users' touch behavior while using their cell phones might be used to authenticate them, which would solve the problem. The performance of the authentication process may be influenced by the attributes chosen (from these behaviors). The purpose of this study is to present an effective authentication technique that implicitly offers a better authentication method for smartphone usage while avoiding the cost of a particular device and considering the constrained capabilities of smartphones. We began by concentrating on feature selection methods utilizing the grey wolf optimization strategy. The random forest classifier is used to evaluate these tactics. The testing findings demonstrated that the grey wolf-based methodology works as a better optimum feature selection for building an implicit authentication mechanism for the smartphone environment when using a public dataset. It achieved a 97.89% accuracy rate while utilizing just 16 of the 53 characteristics like utilizing minimum mobile resources mainly; processing power of the device and memory to validate individuals. Simultaneously, the findings revealed that our approach has a lower equal error rate (EER) of 0.5104, a false acceptance rate (FAR) of 1.00, and a false rejection rate (FRR) of 0.0209 compared to the methods discussed in the literature. These promising results will be used to create a mobile application that enables implicit validation of authorized users yet avoids current identification concerns and requires fewer mobile resources.

Keywords: Smartphone authentication; implicit authentication; grey wolf; random forest; feature selection

1 Introduction

Feature selection is utilized to identify important features from irrelevant features of a predefined feature set [1,2]. The key objectives of feature selection are to reduce data dimensionality and improve prediction performance. In real-world cases, many features often reside in data representation, manifesting as redundant features. This results in a situation where some features take the roles of others. On the other hand, the essential features have a degree of effect on the output, which contains important information that may be reduced if any feature is removed [3].

Identifying the key set of features is a difficult and costly task from a computational point of view. In recent years, metaheuristics have been an effective and more reliable tool for resolving various optimization problems [4]. Metaheuristics have been demonstrated to be more significant in terms of performance when compared to current algorithms, as they do not need to analyze the full search space.

Smartphones are subjected to numerous logins for several reasons, including gaining access to social media, making phone calls, and conducting banking activities. If the authentication procedure is done subliminally, it would be more convenient (and valuable) (without using a password or personal identification number (PIN) or using face recognition or thumbprints). The authentication method implicitly emphasizes adding to the strategies for preventing unauthorized access to mobile devices [5]. This approach operates in the device's background to determine if it should continue or be locked. It is split into two steps. Firstly, the user uses their phone as usual, and the system records their behavior-based characteristics (such as how they tap on the tactile screen). Secondly, when a user gains access to their phone through one of the easy authentication schemes, the method evaluates existing usage patterns against the trained user framework to estimate whether to provide access or lock the phone [6]. Behavioral biometrics is the term for this approach. It analyzes information such as handwriting shape, unique patterns in one's stride, keystroke timing, voice, and other features of one's comprehensive behavior without using additional hardware [7]. Behavioral biometrics offer an advantage over physiological biometrics in that they may be utilized to create a more efficient and constant authentication system. Furthermore, negotiating fingerprints requires no technology to capture behavioral data, making them highly cost-effective.

One new area in this authentication of smartphone usage is the cybersecurity of smartphone activity. A study of 300 smartphone users' reported cybersecurity behaviors and practices was conducted online [8]. Systematic analysis of the respondent data was performed to determine how frequently suggested cybersecurity behaviors and procedures are adopted [9]. The hypotheses were tested using Pearson's chi-square with a 5% significance threshold. Post hoc analysis with Bonferroni correction was carried out for statistically significant connections. To identify and detect such cybersecurity activities, several machine learning and optimization techniques are applied [10].

This work emphasizes developing an effective authentication method for smartphone users that provides implicit authentication without requiring additional hardware and addresses smartphone computing limits. To achieve this goal, we proposed to employ the grey wolf optimizer (GWO) to select the most important feature that can distinguish the authorized user from theft or unauthorized users. The work is evaluated using a public dataset [11], (i.e., touchscreen keystrokes) to increase authentication accuracy while addressing smartphone computing limits (memory size, battery life, and limited hardware). This article's role is to employ the grey wolf as a heuristic search method to identify the most discriminating features for the mobile authentication system. The evaluation and analysis revealed that the suggested technique could overcome the benchmark smartphone authentication systems. The contributions of this paper can be summarized as investigating the effectiveness of

employing the GWO as a feature selection method to develop an implicit authentication approach for smartphone user identification.

The following is a breakdown of the paper's structure. The second section contains related work. The proposed method's methodology is described in detail in the third section. The simulation findings are presented and elaborated on in the fourth section. Finally, the results of the proposed method are summarized and discussed in Section 5.

2 Related Works

In the last few decades, authentication has generally been based on the user's knowledge of a specific thing, referred to as a knowledge-based authentication method. However, research has shown that when this method is applied to mobile devices, it encounters challenges such as low security and a lack of user friendliness [12]. Several studies have been conducted in recent years from this perspective. In this section, related works on mobile authentication will be presented. Furthermore, grey Wolf's capabilities for selecting the best features for mobile authentication will also be presented to emphasize its application to mobile authentication.

In [13], the authors studied how users behave based on their perceptions of safety and convenience. The findings demonstrate that convenience is the most important thing to consider while locking a mobile phone screen. A survey of continuous authentication and behavioral biometrics systems for mobile devices was undertaken by [14]. The authors classify behavioral biometric methods and explain how to authenticate mobile devices. A critical discussion of the literature was provided, along with an outline of the lessons learned and research challenges. With the trend to protect smartphones through the authentication of users, Nader et al. [15] suggested a fusion authentication approach composed of two types of authentication: implicit authentication and continuous authentication. Various features were extracted from users' interactions with Android smartphones for the experiment. The result indicated that neural network classifiers are a better option for authentication for different users. Their findings indicate some promise for mobile user authentication.

A universal system for the evaluation of user authentication methods was proposed in [16]. The system takes close consideration the mobile authentication of users. The proposed system conducts good processing of many users' features. Hence, the system also supported feature processing methods and performance indices to pave the way for enhanced authentication. El-Soud et al. [17] proposed an implicit authentication method for mobile phone authentication. The authors also build the technique, so there will be no additional cost for additional hardware resources. They concluded that the filter-based strategy is the optimal feature extraction for an implied authentication mechanism. Another work presented by Rogoeski et al. [18] conducted a review study on the issues of mobile device authentication. The authors put more emphasis on smartphones and tablets that have many sensors. The work further outlines the limitations of existing traditional user authentication methods and discusses authentication systems that utilize biometric features. Lastly, the authors discuss the potential of using biometrics for the mobile authentication of users.

Researchers have proposed various strategies over the years to maintain privacy for calling or different mobile data for rigid presumptions. However, this can restrict its practical application. The author in [19,20] suggests the reversible data transformation (RDT) algorithm-based data gathering protocol. Their protocol eliminates the need for a private channel and doesn't rely on external authentication to preserve privacy against processing outside its intended use. The release chance of the insider disclosure assault won't be higher than one due to group creation. Similarly, the reversible privacy-preserving data mining strategy safeguards processing that goes above what is required.

Results from the experiment show how beneficial the suggested procedure is and how it may be used in smartphone application recommender systems.

In a study by Wang et al. [21], they surveyed the existing mobile authentication methods. The approaches, threats, and trends in the field of study were outlined in detail. Another study by Karakaya et al. [22] used sensor data from the user's hand movement orientation and grasp (HMOG) for smart device authentication. The authors used four machine-learning techniques: decision forest, boosted decision tree, support vector machine (SVM), and logistic regression. The decision forest algorithm gives the best result based on the obtained results.

Various grey wolf optimizer-based feature extraction techniques were presented for classification and feature selection. Reviews were also conducted on the methods and their applications [23,24]. For feature selection, the author proposed combining hybrid binary grey wolf optimization with particle swarm optimization [BGWOPSO] [25]. Based on an experiment conducted, the results show that BGWOPSO significantly outperformed other techniques, such as the binary grey wolf optimization (BGWO), the binary particle swarm optimization (PSO), and the binary genetic algorithm. In another work [26], the authors proposed an improved binary grey wolf optimizer (IGWO) for the feature selection method. Their findings demonstrate that the IGWO algorithm has global search capabilities and high efficiency, making it suitable for reliability analysis in engineering. Chantar et al. [27] proposed an enhanced binary grey wolf optimizer. They investigated its performance employing different machine learning methods such as decision trees (DT), *K*-nearest neighbor (KNN), Naive Bayes (NB), and SVM classifiers.

The results show some enhancement from using the grey wolf optimizer for selecting features with the SVM classifier. To choose the best features for detecting coronary artery disease, the author in [28] suggested a feature selection approach based on grey wolf optimization and SVM. The result shows some promise in the proposed method compared to existing practices. In [29], the authors suggested a binary grey wolf optimization technique. Their work aims to extract the most appropriate features from a biomedical dataset. The result proves the capability of their approach and the importance of using grey wolf optimization for feature extraction. Tahoun et al. [4] used the grey wolf optimizer to suggest a technique for extracting features from wavelet and curvelet sub-bands for mammogram classification. The result shows that the best features are effectively extracted when binary grey wolf optimization is utilized.

Lastly, Salih et al. [30] proposed a model for improving the performance of network intrusion detection systems. The author used an Anaconda Python open-source platform to deploy a set of algorithms for selecting features for authentication. An evaluation was conducted on the collected features using a deep learning approach. The result reveals that an intrusion detection system's accuracy has improved.

From the literature review in this section, we have retrieved many remarks that need to be considered. We observed that there are various mobile device authentication methods proposed, with more emphasis on using machine learning algorithms and biometric features for user authentication. Their limitations were also explored by the studies. In addition, various advancements in this domain have been recorded. Furthermore, works on grey wolf optimization for feature extraction were also highlighted. As a result, using grey wolf optimizers for feature extraction has proven to be beneficial. To the author's knowledge, it has never been used in the research domain of mobile authentication. Therefore, this article aims to utilize the grey wolf optimization method for feature extraction to help authenticate users on mobile devices.

3 Preliminaries

This section provides details of the different approaches and algorithms that are used in the suggested method. It focuses on the grey wolf optimization (GWO) technique that can be utilized for selecting features. It also detailed the classifier utilized, namely the random forest classifier.

3.1 Feature Selection: Binary Grey Wolf Optimizer

In [31], the authors suggested the use of a metaheuristic optimization approach called GWO. This name is given based on the algorithm's similarity to how grey wolves hunt and form the leadership hierarchy. In GWO, the size of the pack is usually restricted between five and twelve. Moreover, the GWO's population is divided into different tiers, including Omega, Alpha, Delta, and Beta levels, where the Alpha status is given to the leader of the pack. The Alpha handles all the pack's critical decisions. Moving down the hierarchy, the wolf belonging to the Beta level is responsible for assisting the Alpha wolf in various tasks. Below them comes the Delta category, which provides security to the pack and follows the orders given to it by the Alpha and Beta wolves. Furthermore, they are superior to Omega category wolves. These are the lowest in the hierarchy and are responsible for performing scouting [32]. They are essential in supplying and conducting reconnaissance, warning of danger, and guarding the pack against outsiders. The GWO provided in [33] was employed in this research to identify the most suitable set of features that can identify the correct user.

3.2 Classification: Random Forests

Random forests [34] is an ensemble learning algorithm for supervised machine learning. The random forest ensemble notion states that merging numerous weak classifiers can result in an accurate classification rate. To improve prediction accuracy, the random forest combined the results of multiple decision trees. To assess the accuracy of the suggested strategy, the classification phase in this study was done using 10-fold cross-validation. The random forest model is illustrated in Fig. 1.

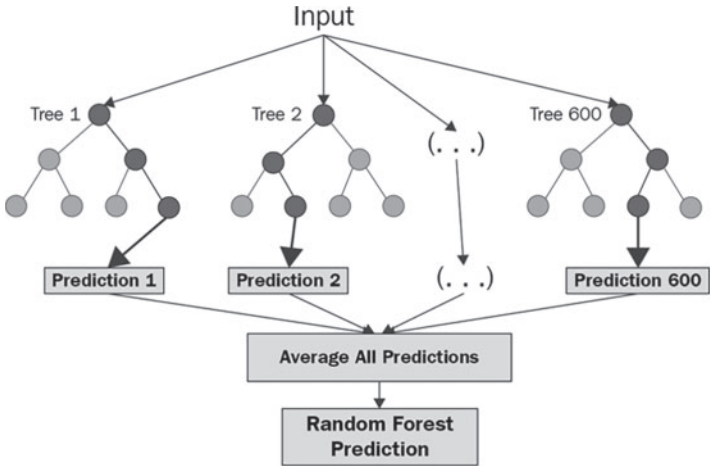

Figure 1: Random forest working model

4 Proposed System

The suggested system, represented in Fig. 2, is divided into three major phases: feature selection, sampling of features, and classification. The GWO algorithm [35] ranks the features during the feature

selection phase. This phase produces ranked feature sets as a result. The hypothesis is that these ranked features will produce different classification results depending on the optimization approach used to sort the features.

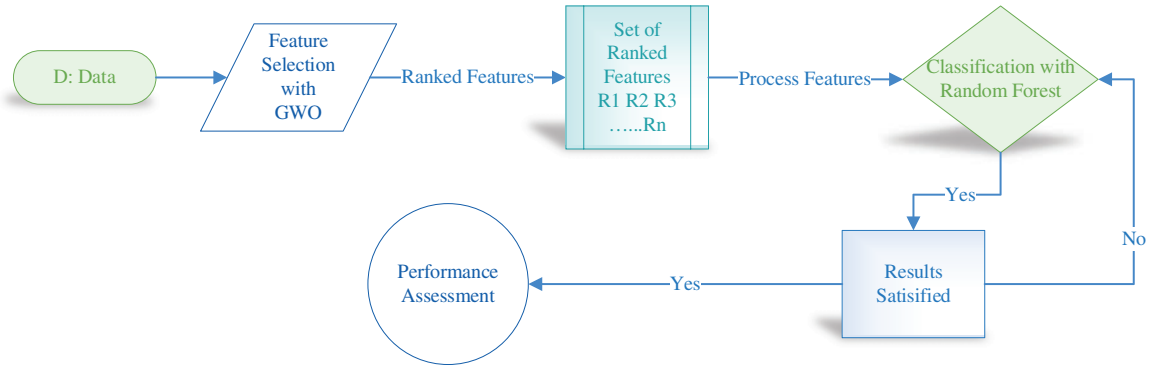

Figure 2: The proposed system

4.1 Data Set Description

In this study, the public benchmark data set namely Rafik Hariri University (RHU) touch mobile keystroke was used [11]. The data set was collected from 51 individuals who were asked to input the password "rhu.university" 15 times in 3 meetings with a normal interval of 5 days between sessions. The data is collected independently in each session. The database obtained comprises 954 samples, including males and females of various ages. The dataset collects the time between two key presses (PP), the time between applying pressure to the key and releasing the key (PR), the interval between two key releases (RR), and the time interval between releasing the key and pressing on the key (RP) from each user [36]. The fundamental characteristics (PP, PR, RR, RP) are augmented by sub-features that define the different time intervals for inputting a password: PP, RR, and RP each contain 13 sub-features, while PR includes 14 instances of sub-features. As a result, each user's total number of characteristics is 53 (i.e., classes). Table 1 shows the feature numbers corresponding to its notation.

Table 1: The features numbers corresponding to their notation

Feature	Denoted	Feature	Denoted	Feature	Denoted	Feature	Denoted
F1	PP_1	F14	PR_1	F28	RR_1	F41	RP_1
F2	PP_2	F15	PR_2	F29	RR_2	F42	RP_2
F3	PP_3	F16	PR_3	F30	RR_3	F43	RP_3
F4	PP_4	F17	PR_4	F31	RR_4	F44	RP_4
F5	PP_5	F18	PR_5	F32	RR_5	F45	RP_5
F6	PP_6	F19	PR_6	F33	RR_6	F46	RP_6
F7	PP_7	F20	PR_7	F34	RR_7	F47	RP_7
F8	PP_8	F21	PR_8	F35	RR_8	F48	RP_8
F9	PP_9	F22	PR_9	F36	RR_9	F49	RP_9
F10	PP_10	F23	PR_10	F37	RR_10	F50	RP_10
F11	PP_11	F24	PR_11	F38	RR_11	F51	RP_11

(Continued)

Table 1: Continued

Feature	Denoted	Feature	Denoted	Feature	Denoted	Feature	Denoted
F12	PP_12	F25	PR_12	F49	RR_12	F52	RP_12
F13	PP_13	F26	PR_13	F40	RR_13	F53	RP_13
		F27	PR_14				

5 Experimental Results

The experiments are performed using the MATLAB simulator, and the hardware specifications are 8 GB of RAM and a Core i7 (Intel CPU) with a dual 3.40 GHz processor. The outcomes are assessed using a number of well-known user authentication mechanism evaluation measures, including EER, FAR, accuracy, and FRR. All these metrics are derived using 10-cross fold, a validation method that implies that each sample of data will appear exactly once in testing samples and nine times in training samples (10-1). The 10-cross-fold scheme was applied to guarantee that the statistical analysis outcomes could also be applied to other datasets.

Table 2 summarizes the outcomes of the suggested technique. This table shows that the most significant results (97.8947 percent) were obtained when the number of features was set to 16. Table 3 shows a list of the main features chosen using the grey wolf selection approach. It can be determined that the random forest can identify the proper person using just 16 out of 53 features with a 1.00 FAR, a 0.0209 FRR, 97.8947 percent accuracy, and a 0.05104 ERR. Furthermore, the results achieved are superior in terms of throughput, consuming only 44.319118 milliseconds. Based on the sub-features identified in Table 3, it is possible to conclude that the grey wolf method, along with the random forest approach, might be utilized to develop an effective underlying authentication technique for new-era smartphone apps, although satisfying its limited computing performance.

Table 2: The obtained results from a distinct set of features

Input	No of features	Accuracy	FAR	FRR	EER	Time
All features	53	93.1579	3.7500	0.1457	1.9478	91.942723
GW1	22	95.6989	4.0000	0.0881	2.0441	56.021934
GW2	23	94.8454	5.0000	0.1066	2.5533	57.132486
GW3	25	95.7895	3.5000	0.0862	2.0431	58.588774
GW4	22	95.8333	3.5000	0.0853	1.7926	54.360930
GW5	16	97.8947	1.0000	0.0209	0.5104	44.319118
GW6	23	93.8776	5.5000	0.1279	2.8139	55.129522
GW7	29	95.6522	3.0000	0.0891	1.5446	63.597622
GW8	21	93.6170	6.5000	0.1824	3.3412	52.270956
GW9	25	95.8333	3.5000	0.0853	1.7926	58.906093
GW10	21	68.63	3.4713	0.0546	1.7630	138.886755
GW11	24	91.4894	5.8333	0.1824	3.0079	57.562452
GW12	20	94.7368	5.0000	0.1089	2.5545	50.907952
GW13	21	45.1613	33.5000	0.0488	16.7744	145.689397
GW14	18	43.3628	39.0000	0.0525	19.5263	138.474307

(Continued)

Table 2: Continued

Input	No of features	Accuracy	FAR	FRR	EER	Time
GW15	22	46.3158	34.4167	0.0466	17.2316	143.954861
GW16	24	37.8947	33.1667	0.0659	16.6163	140.794354
GW17	23	43.6170	36.5000	0.0520	18.2760	140.546223
GW18	21	41.0526	35.5000	0.0578	17.7789	139.672929
GW19	21	48.3871	30.8000	0.0429	15.4215	139.439267
GW20	24	46.8085	33.6667	0.0457	16.8562	141.432471

Also, the accuracy has been improved using a few features; these results support the claim that more features might struggle with the classifier and reduce the accuracy of the classification system. At the same time, the accuracy started to decrease with the use of the set of features obtained from GW13. This can be interpreted to mean that each collection of features could produce a high accuracy rate, while others might reduce the classification rate. In other words, grouping the features is particularly important in identifying the most discriminant set of features. This might mean that the high-impact features contribute significantly to the user authentication procedure (the identified set of GW5). When low-importance characteristics were used, the results were reduced because these elements did not influence on differentiating distinct consumers. Furthermore, the processing time has grown, the accuracy rate has decreased, and the error rate has increased.

Table 3: The selected features for each grey wolf run

Input	Selected features									
GW1	F5	F25	F24	F21	F19	F17	F16	F15	F14	F11
	F26	F52	F49	F48	F43	F37	F35	F33	F29	F28
GW2	F2	F5	F24	F22	F21	F20	F19	F18	F15	F14
	F25	F26	F42	F40	F39	F36	F33	F32	F29	F28
	F43	F47	F53	F51	F50	F48				
GW3	F8	F26	F25	F24	F19	F18	F15	F14	F13	F12
	F27	F47	F41	F40	F38	F37	F36	F32	F30	F29
	F49									
GW4	F5	F8	F6	F20	F19	F18	F17	F16	F14	F13
	F23	F29	F27	F41	F40	F38	F37	F36	F34	F32
	F43	F50	F49	F22	F21	F19	F18	F16	F53	F52
GW5	F6	F10	F9	F32	F28	F27	F25	F24	F13	F12
	F35	F37	F36	F51	F49	F47	F43	F40		
GW6	F6	F8	F23	F23	F20	F19	F16	F15	F13	F10
	F24	F25	F42	F43	F39	F38	F36	F33	F28	F27
	F49	F50	F52							
GW7	F1	F2	F17	F19	F15	F13	F10	F9	F8	F3
	F21	F22	F23	F24	F39	F38	F29	F25	F24	F23

(Continued)

Table 3: Continued

Input	Selected features									
GW8	F25	F29	F52	F48	F45	F41	F39	F38		
	F3	F5	F17	F19	F16	F14	F13	F12	F9	F8
	F21	F22	F39	F40	F37	F36	F29	F26	F25	F24
	F41	F48	F51	F50	F49					
GW9	F1	F2	F19	F20	F17	F15	F13	F11	F9	F8
	F21	F22	F30	F32	F29	F28	F27	F25	F24	F23
	F33	F37	F53	F52	F49	F46	F45	F42	F40	
GW10	F1	F5	F22	F24	F20	F19	F18	F17	F13	F9
	F25	F26	F42	F45	F41	F37	F36	35	F30	F27
	F49	F51	F53	F52						
GW11	F5	F12	F25	F28	F23	F21	F19	F16	F14	F13
	F32	F35	F48	F50	F43	F42	F40	F38	F37	F36
	F53									
GW12	F8	F9	F23	F25	F20	F19	F18	F16	F13	F10
	F28	F29	F49	F51	F42	F39	F36	F35	F32	F30
	F53									
GW13	F2	F7	F19	F20	F17	F16	F15	F12	F11	F9
	F23	F24	F40	F43	F37	F36	F35	F33	F32	F28
	F50	F53								
GW14	F1	F2	F16	F18	F14	F13	F10	F8	F5	F3
	F21	F22	F40	F41	F37	F36	F35	27	F24	F23
	F51	F52	F53							
GW15	F5	F6	F22	F24	F21	F20	F19	F11	F10	F8
	F26	F27								
GW16	F8	F9	F20	F22	F19	F17	F16	F15	F14	F13
	F24	F26	F41	F42	F40	F39	F37	F36	F33	F27
	F43	F45	F53	F52	F51					
GW17	F1	F2	F17	F20	F15	F14	F13	F10	F8	F6
	F21	F24	F40	F41	F39	F36	F33	F32	F26	F25
	F43	F50	F53	F51						
GW18	F9	F13	F23	F24	F21	F20	F19	F18	F16	F14
	F25	F26	F40	F45	F36	F35	F33	F30	F29	F28
	F46	F50	F53	F52	F51					
GW19	F2	F5	F20	F22	F16	F13	F12	F10	F9	F6
	F25	F27	F43	F45	F42	F41	F40	F37	F36	F35
	F47	F48	F51							
GW20	F1	F2	F23	F25	F20	F19	F18	F15	F9	F8
	F27	F32	F41	F43	F40	F39	F38	F37	F36	F35
	F49	F53								

To further assess our findings, we compared them to the relevant study in Section 2. The comparative works were chosen so that the suggested authentication implicitly applies to cell phones via either freely available or private datasets. Table 4 provides an overview of this comparison. The following observations may be derived from this table: To begin, using the private dataset, [5] attained a significant accuracy of 92.1 percent. This study employed private data, which is not publicly available for the community to review. Furthermore, neither [6] nor [7] reported any data for FRR, ERR, or FAR, which are critical when evaluating any authentication technique.

Table 4: Presents a comparison between the related work against our proposed method

Reference	Classifier	Feature selection	Acc. (%)	Dataset	FAR	No. of users	FRR	No. of features	ERR
[5]	k-NN	N/A	92.1	Private	7.5	20	8.3	N/A	N/A
[6]	K-NN	t-test	91.00	Public	N/A	51	N/A	47	N/A
[7]	Bagging	GA	83.80	Public	N/A	51	N/A	10	N/A
[14]	RF	Rank	97.80	Public	2.03	51	0.04	25	1.04
Our method	RF	Grey wolf	97.8947	Public	1.0000	51	0.0209	16	0.5104

Second, when compared to previous schemes that utilized a similar dataset (i.e., [11]) and employed feature selection strategies (i.e., [6,7]), they utilized just ten characteristics (compared to our method's 16 characteristics) and achieved an 83.8 percent accuracy, which is around 14% less than ours. Furthermore, the performance of the approach [7] is not assessed with evaluation metrics, i.e., ERR, FAR, or FRR. While the approach in [14] used 25 features as input and produced an acceptable result, our method outperforms it in terms of accuracy, the number of features, FAR, FRR, and ERR. Furthermore, the time complexity of our proposed model is 13% less than that of the existing methods.

In conclusion, this study reveals that our proposed method is viable in terms of computational cost (effectiveness), quantifying the number of resources necessary to accomplish implicit authentication. As previously said, implicit authentication is a constant procedure in the smartphone's background. As a result, it would be ideal if such an authentication system used as few smartphone resources as possible, such as computing power and memory. The next generation of smartphones might be said to improve mobile CPUs and memory. However, the cap is due to space and heat transfer limits. As a result, it is projected that battery life will be a significant hurdle to mobile computing efficiency in the next few years. As a result, mobile authentication systems that consume as minimal energy as possible would be practicable. Our proposed technique performed implicit user authentication with only 16 out of 53 characteristics, indicating that it is efficient. This will also reduce the server's computation cost (using 25 features as an alternative to 53). In the paradigm of the thin-client procedure, our proposed solution can be utilized for various uses, including banks, the security sector, government administration, and healthcare. With more usage, an additional protection cover is necessary to secure the customers' personal information.

Regarding user authentication security, the comparison provided in Table 4 demonstrates that our solution obtained lower FRR, FAR, and ERR rates, which are significant for avoiding giving illicit access to users' smartphones and storing private information like banking and personal information.

6 Conclusion

With the spread of smartphone devices, an easy technique to authenticate their users will emerge. To address this issue, our study presented an improved implicit authentication approach. The proposed approach combines the grey wolf algorithm as a feature selection approach with a random forest classifier to automatically authenticate the user based on their touch behavior. The most significant characteristics are picked and fed into the random forest classifier to establish which user is accessing the smartphone. The evaluation results revealed that a smartphone operator might be indirectly authenticated using fewer attributes (16 out of 53) selected using the grey wolf optimizer technique and classified by the random forest while achieving a reduced error rate: 0.0209 FRR, 0.5104 ERR, and 1.00 FAR. The smartphone's limits (such as processing capability, battery life, and memory size) may be addressed by fewer features. Furthermore, shoulder surfing and security assaults might be stopped. Future research will examine whether deep learning approaches may enhance accuracy and other measures.

Acknowledgement: This work was funded by the University of Jeddah, Jeddah, Saudi Arabia, under grant No. (UJ-21-DR-25). The authors, therefore, acknowledge with thanks the University of Jeddah technical and financial support.

Funding Statement: This work was funded by the University of Jeddah, Jeddah, Saudi Arabia, under grant No. (UJ-21-DR-25). The authors, therefore, acknowledge with thanks the University of Jeddah technical and financial support.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. Hamed and H. Nassar, "Efficient feature selection for inconsistent heterogeneous information systems based on a grey wolf optimizer and rough set theory," *Soft Computing*, vol. 25, no. 24, pp. 15115–15130, 2021.

[2] D. M. Ablel-Rheem, A. O. Ibrahim, S. Kasim, A. A. Almazroi and M. A. Ismail, "Hybrid feature selection and ensemble learning method for spam email classification," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 1.4, pp. 217– 223, 2020.

[3] O. Almomani, "A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms," *Symmetry*, vol. 12, no. 6, pp. 1046, 2020.

[4] M. Tahoun, A. A. Almazroi, M. A. Alqarni, T. Gaber, E. E. Mahmoud *et al.,* "A grey wolf-based method for mammographic mass classification," *Applied Sciences*, vol. 10, no. 23, pp. 8422, 2020.

[5] W. H. Lee and R. Lee, "Implicit sensor-based authentication of smartphone users with smartwatch," in *Proc. of the Hardware and Architectural Support for Security and Privacy*, Seoul, Republic of Korea, pp. 1–8, 2016.

[6] W. H. Lee and R. B. Lee, "Multi-sensor authentication to improve smartphone security," in *Int. Conf. on Information Systems Security and Privacy (ICISSP)*, Angers, France, pp. 1–11, 2015.

[7] A. Tharwat, A. Ibrahim, T. Gaber and A. E. Hassanien, "Personal identification based on mobile-based keystroke dynamics," in *Proc. of the Int. Conf. on Advanced Intelligent Systems and Informatics*, Cairo, Egypt, pp. 457–466, 2018.

[8] K. Shaukat, S. Luo, V. Varadharajan, I. A. Hameed, S. Chen *et al.,* "Performance comparison and current challenges of using machine learning techniques in cybersecurity," *Energies*, vol. 13, no. 10, pp. 2509, 2020.

[9] K. Shaukat, S. Luo, V. Varadharajan, I. A. Hameed and M. Xu, "A survey on machine learning techniques for cyber security in the last decade," *IEEE Access*, vol. 8, no. 1, pp. 222310–222354, 2020.

[10] K. Shaukat, S. Luo, S. Chen and D. Liu, "Cyber threat detection using machine learning techniques: A performance evaluation perspective," in *2020 Int. Conf. on Cyber Warfare and Security (ICCWS)*, Virginia, USA, IEEE, pp. 1–6, 2020.
[11] M. El-Abed, M. Dafer and R. E. Khayat, "RHU keystroke: A mobile-based benchmark for keystroke dynamics systems," in *Int. Carnahan Conf. on Security Technology (ICCST)*, Rome, Italy, pp. 1–4, 2014.
[12] I. Stylios, S. Kokolakis, O. Thanou and S. Chatzis, "Behavioral biometrics and continuous user authentication on mobile devices: A survey," *Information Fusion*, vol. 66, no. 2, pp. 76–99, 2021.
[13] P. K. Sari, G. S. Ratnasari and A. Prasetio, "An evaluation of authentication methods for smartphone based on users' preferences," *IOP Conference Series: Materials Science and Engineering*, vol. 128, no. 1, pp. 12036, 2016.
[14] I. C. Stylios, O. Thanou, I. Androulidakis and E. Zaitseva, "A review of continuous authentication using behavioral biometrics," in *Proc. of the South East European Design Automation, Computer Engineering, Computer Networks and Social Media Conf.*, Kastoria, Greece, pp. 72–79, 2016.
[15] J. Nader, A. Alsadoon, P. W. C. Prasad, A. K. Singh and A. Elchouemi, "Designing touch-based hybrid authentication method for smartphones," *Procedia Computer Science*, vol. 70, pp. 198–204, 2015.
[16] D. Progonov, V. Prokhorchuk and A. Oliynyk, "Evaluation system for user authentication methods on mobile devices," in *11th Int. Conf. on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine, IEEE, pp. 95–101, 2020.
[17] M. W. A. El-Soud, T. Gaber, F. AlFayez and M. M. Eltoukhy, "Implicit authentication method for smartphone users based on rank aggregation and random forest," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 273–283, 2021.
[18] M. Rogowski, K. Saeed, M. Rybnik, M. Tabedzki and M. Adamski, "User authentication for mobile devices," in *IFIP Int. Conf. on Computer Information Systems and Industrial Management*, Krakow, Poland, pp. 47–58, 2013.
[19] U. Javed, K. Shaukat, I. A. Hameed, F. Iqbal, T. M. Alam *et al.,* "A review of content-based and context-based recommendation systems," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 3, pp. 274–306, 2021.
[20] K. Shaukat, F. Iqbal, T. M. Alam, G. K. Aujla, L. Devnath *et al.,* "The impact of artificial intelligence and robotics on the future employment opportunities," *Trends in Computer Science and Information Technology*, vol. 5, no. 1, pp. 050–054, 2022.
[21] C. Wang, Y. Wang, Y. Chen, H. Liu and J. Liu, "User authentication on mobile devices: Approaches, threats and trends," *Computer Networks*, vol. 170, no. 2, pp. 107–118, 2020.
[22] N. Karakaya, G. I. Alptekin and Ö. D. İncel, "Using behavioral biometric sensors of mobile phones for user authentication," *Procedia Computer Science*, vol. 159, no. 5, pp. 475–484, 2019.
[23] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, S. Mirjalili and H. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," in *Evolutionary Machine Learning Techniques, Algorithms for Intelligent Systems*, Springer, Singapore, pp. 273–286, 2020. https://doi.org/10.1007/978-981-32-9990-0_13
[24] H. Faris, I. Aljarah, M. A. Al-Betar and S. Mirjalili, "Grey wolf optimizer: A review of recent variants and applications," *Neural Computing and Applications*, vol. 30, no. 2, pp. 413–435, 2018.
[25] Q. Al-Tashi, S. J. A. Kadir, H. M. Rais, S. Mirjalili and H. Alhussian, "Binary optimization using hybrid grey wolf optimization for feature selection," *IEEE Access*, vol. 7, pp. 39496–39508, 2019.
[26] A. Hraiba, A. Touil and A. Mousrij, "Improved grey-wolf optimizer for reliability analysis," in *Int. Conf. on Advanced Intelligent Systems for Sustainable Development*, Marrakech, Morocco, pp. 88–98, 2019.
[27] H. Chantar, M. Mafarja, H. Alsawalqah, A. A. Heidari, I. Aljarah *et al.,* "Feature selection using binary grey wolf optimizer with elite-based crossover for Arabic text classification," *Neural Computing and Applications*, vol. 32, no. 16, pp. 12201–12220, 2020.
[28] Q. Al-Tashi, H. Rais and S. Jadid, "Feature selection method based on grey wolf optimization for coronary artery disease classification," in *Int. Conf. of Reliable Information and Communication Technology*, Kuala Lumpur, Malaysia, pp. 257–266, 2018.

[29] E. Emary, H. M. Zawbaa and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, no. 1, pp. 371–381, 2016.
[30] A. A. Salih, S. Y. Ameen, S. R. Zeebaree, M. A. Sadeeq, S. F. Kak *et al.,* "Deep learning approaches for intrusion detection," *Asian Journal of Research in Computer Science*, vol. 9, no. 4, pp. 50–64, 2021.
[31] S. Mirjalili, S. M. Mirjalili and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
[32] H. Faris, I. Aljarah, M. A. Al-Betar and S. Mirjalili, "Grey wolf optimizer: A review of recent variants and applications," *Neural Computing and Applications*, vol. 30, no. 2, pp. 413–435, 2018.
[33] J. Too and A. R. Abdullah, "Opposition based competitive grey wolf optimizer for EMG feature selection," *Evolutionary Intelligence*, vol. 14, no. 4, pp. 1691–1705, 2021.
[34] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintha *et al.,* "Improved random forest for classification," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4012–4024, 2018.
[35] S. Mirjalili, "How effective is the grey wolf optimizer in training multi-layer perceptrons," *Applied Intelligence*, vol. 43, no. 1, pp. 150–161, 2015.
[36] H. Zhang, C. Yan, P. Zhao and M. Wang, "Model construction and authentication algorithm of virtual keystroke dynamics for smart phone users," in *IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, Budapest, Hungary, pp. 000171–000175, 2016.