Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 4, pp. 1919–1941, DOI 10.12694/scpe.v26i4.4631

ENHANCED CRIMINAL SUSPECT IDENTIFICATION USING A NOVEL SMART SYSTEM WITH HYBRID ENCRYPTION AND ANN CLASSIFICATION

NAJAH KALIFAH ALMAZMOMI[*]

Abstract. The current advancement in the growth of crime rates and the complexity of crimes that are being committed in society call for the adoption of better methods of identifying criminal suspects. While traditional methods are somewhat useful, they are not sufficient in dealing with the intricacies and the sheer scale of today's data landscape. Machine intelligence has become popular in many fields including criminal justice because of its capability to learn from big data. This study presents a novel smart system for classifying criminal suspects using four key steps: query-based authentication (QBA), data categorization, data encryption and decryption, and artificial neural network (ANN)-based classification. QBA ensures only authorized access to sensitive data by verifying user-specific information. Data is categorized into sensitive (personal and social criminal data) and non-sensitive (classification results) categories, with sensitive data encrypted using a Two-Level hybrid ECC (Elliptic Curve Cryptography) and ECC-RSA (Rivest, Shamir, Adleman) model, optimized via the HSMEO algorithm for high security and efficiency. The ECC-RSA model outperforms traditional encryption methods (AES, DES, RSA, ECC) in security (98.53 %), trust score (4.83), throughput (77.57), and encryption/decryption times (3.459/2.994 seconds). Additionally, the HSMEO model significantly reduces key generation time to 1.97545 seconds, surpassing other optimization strategies like SMO, EO, PSO, MFO, and FFO. Graphical representations of key metrics validate the ECC-RSA model's superior performance in security, efficiency, and reliability, making it an effective method for protecting sensitive data and ensuring efficient criminal suspect classification through expert system.

Key words: Criminal Suspect Classification; Machine Learning; ECC-RSA; HSMEO; ANN.

1. Introduction.

In the current digital era, the precarious task of the proliferation of criminality in online social networks is a real threat to the planning and implementation of illegal operations. The very essence of these platforms that allow automatic interaction among a wide number of people coupled with anonymity has unintentionally created a hospitable area for mischievous activities from cyberbullying to international crime [1]. Despite the magnitudes of the problem, the field of studies is severely underrepresented in terms of the amount of research carried out that focuses on identifying and categorizing perpetrators of cyberbullying. This lack of scholarly investigation, in turn, has emphasized the urgent necessity of tapping into cutting-edge technologies, such as machine intelligence, to overcome the menace of digital untruths [2].

The foregoing discussion underscored the backdrop of the current research that is set to expand the frontiers of machine intelligence by developing smart system algorithms that can be used to identify and categorize individuals involved in online criminal activities [3]. This research strategy strongly relies on the data derived from these platforms besides the skills and knowledge gained from the law enforcement databases and the ultimate goal is to create a robust expert system framework that is used in identifying possible criminals. Essentially, this research proposal calls for a change of paradigm where social networking and law enforcement are brought on the same footing to address the current growing cases of digital insecurity [4].

The rapid development of the criminal activities that are performed through online social networks has turned into a huge problem for the new generation, which is on the one hand caused by the unique opportunity that cyber-criminals are given to harmonize their actions and to implement their illegal schemes [5]. Although the weight of this matter is immense, nonetheless, the state of the research about who committed those digital misdeeds and the way they can be classified lacks depth. This research gap points out the importance of new technologies like machine intelligence that can help us identify and categorize suspects among the many social media interactions this complex network involves [6].

With the development of machine intelligence and its ability to solve a complex pattern and anomalies

[*]Department of MIS, College of Business, University of Jeddah, Jeddah, Kingdom of Saudi Arabia (nalmazmomi@uj.edu.sa)

in much bigger data, the present study focuses on the utilization of ANNs to meet this challenge [7]. Such a research project aims to enlist social media and law enforcement agencies' big data to develop a more smart system for mapping criminal suspects within online social networks. This venture is determined by the urgency to develop digital security and to take necessary steps against the massive dissemination of cybercrimes in cyberspace [8].

The main parts of the model being the four crucial elements that interplay one with the other ensuring the overall effectiveness and robustness are the pivotal points in its functionality. First, authentication is the initial task carried out by the query protocols, providing the tools needed to access and analyze the data streams [9]. Additionally, the data category is a significant aspect in discriminating sensitive and non-sensitive information, and consequently, the basis for the proceeding analyses [10]. Particularly, the implementation of the ECC-RSA [11] scheme— a next-generation encryption standard that combines RSA and ECC algorithms [12] — has demonstrated the integrity and confidentiality of the data as the key priorities.

Additionally, the employment of HSMEO [13] which is a new kind of method is an innovative step that is used to optimize the private keys choice within the encryption system. The proposed model, with slime mold-inspired algorithmic solutions that have inbuilt efficacy and adaptability, shows promise to increase the degree of both data security and encryption efficiency. Therefore, such solutions would strengthen the integrity of smart system. Finally, the integration of artificial neural networks (ANNs) [14] for criminal suspect classification represents the apex of the smart system and enables it to detect patterns and norm violations within the abundant social network data. This expert system approach improves the model's performance in terms of classification and prediction of criminal behaviour.

This research aims to create a useable taxonomy for the extraction of suspicious individuals from social networks with the help of machine learning algorithms and the neural network classifier (ANN) in particular. Although it is self-evident that there is no alternative to identifying cyber criminals in the social media realm, there is a gap in literature on the effective categorization methodologies. Therefore, the study is aimed to harmonize these two types of data obtained from the web and law enforcement agencies for the identification of possible suspects. This undertaking embraces a smart system that incorporates authentication, data classification, encryption, as well as expert system components like artificial intelligence neural networks to precisely and accurately distinguish and group actual criminals present in social media platforms. The scientific validation of the model is conducted through experimental trials carried out within the Matlab environment, and the model performance indicators are sought. In addition, the study intends to assess the security and efficiency level of the built encryption framework by using the ECC-RSA paradigm and HSMEO technique to generate the private keys. The major contributions of this study are as follows:

- The study presents a unique smart system methodology for classifying criminal suspects, integrating four key steps: query-based authentication (QBA), data categorization, data encryption and decryption, and artificial neural network (ANN)-based classification.
- By implementing QBA, the methodology ensures that only authorized individuals can access sensitive data, enhancing the overall security of the system.
- The approach categorizes data into sensitive and non-sensitive groups, ensuring that personal and social criminal data are securely managed while classification results are efficiently handled.
- The study introduces a Two-Level hybrid ECC and RSA (ECC-RSA) model, optimized with the HSMEO algorithm, providing superior security (98.53%), trust score (4.83), and throughput (77.57) compared to traditional encryption methods.
- The HSMEO algorithm significantly reduces key generation time to 1.97545 seconds, outperforming other optimization strategies such as SMO, EO, PSO, MFO, and FFO, enhancing the efficiency of cryptographic operations.
- Graphical representations of key metrics validate the ECC-RSA model's superior security, efficiency, and reliability performance, establishing it as an effective expert system for protecting sensitive data and efficiently classifying criminal suspects.

The paper's organization is structured as follows: Section 2 provides a comprehensive literature review, examining existing methodologies and highlighting the need for an advanced approach to criminal suspect classification. Section 3 details the material and methods used in the study, describing the novel four-step methodology which includes QBA, data categorization, a Two-Level hybrid ECC and RSA (ECC-RSA) encryption model optimized with the HSMEO algorithm, and ANN-based classification. Section 4 presents the experimental results, showcasing the proposed ECC-RSA model's performance metrics compared to traditional encryption methods, and highlighting its superior security, efficiency, and reliability. Section 5 discusses the findings in-depth, analyzing the implications of the results and how the proposed methodology enhances data protection and classification efficiency. Section 6 concludes the paper by summarizing the key contributions and potential future research directions in criminal suspect classification and data encryption.

2. Literature Review. The paper is going to dwell on the different strategies that are being used to improve the systems of collecting evidence and predicting the suspects within the criminal networks. The work of Jhee et al. [15] is the application of machine learning in criminal investigations. To be specific, the authors developed a fast inference approach for the analysis of large criminal networks. The way they experimented is representative of their precision and fast inference ability, which makes them a good choice for real-life cases. Nevertheless, the test demonstrated that with bigger datasets, the inference becomes slower and it is more complicated to identify data from criminal networks. This shows that more research needs to be done to assess the technology application across diverse crime data. Jhee et al. [16] study also brings forth the criminal network-based suspect prediction framework and the impact of the reduction in execution time while not compromising the competitive performance of the framework is highlighted. Nevertheless, a few obstacles such as the slow running of the entire processing on large datasets and the memory-intensive computations that come with it are the outstanding disadvantages that restrict it from wider adoption. Hence, the investigations of scalable solutions for wider applicability are highly encouraged.

Furthermore, Chachoo [17] puts forward the network analysis of digital crimes that the ANOS OCL method is used for identifying the key actors in criminal networks, paying attention to online social networks. While efficient in detecting criminal activities, it could not fully reflect the roles of all key actors and complex interactions in the crime community. Hence, it is evident that online criminality is more complicated than it seems. Besides that, Gupta et al. [18] provide mechanisms for identifying cybercriminals on social media using machine learning techniques and also, they highlight the difficulties associated with the text format of crime-related activities.

Florentino et al. [19] provide an approach with a novel characteristic: they do not label data anymore; they also do not relate messages to any vocabulary. However, data scarcity and the dynamic nature of social networks remain the major problems that make it imperative to do frequent model updates to be effective. Shafi et al. [20] proposed a reduced-complexity method for classifying criminal activists on social media, where the feature extraction process is simplified and the technique can be adapted to the regular changes in the strategies of criminals.

Alebouyeh and Bidgoly [21] put a spotlight on the criminality of social network users through centrality measures, giving more importance to relationships than the content that is published. Although there is no denying that strong relationships between criminals can happen, yet there is a possibility of centrality measures not being able to detect such kinds of criminals correctly. Adding up to that, ethical issues and hidden biases of centrality-based criminal apprehension should be also examined. On the other hand, Deepak and his colleagues [22] explored the knowledge-based hybridized approach for crime classification by presenting the dynamic ontology generation together with deep learning techniques. The technique shows better results on various datasets, which proves that the creation and use of modern neural networks open new perspectives on crime classification. The paper, though, does not seem to be addressed to limit specific issues, hence the need to carry out additional studies for the comprehensive evaluation of its applicability. These researches, as a whole, pave the way for the development of methods used for the detection of crime and the identification of the suspects within the networks as well as pointing out both the pros and cons of that network in real-life situations.

Florentino, Goldschmidt, and Cavalcanti [23] are offering a vocabulary-driven technique that is supposed to help the identification of people suspected of criminal activity on social networks, which is a difficult task when there is a lack of training datasets. They applied a controlled vocabulary including categorized terms to social media messages which are labeled data, so that it could be analyzed without them. Thru the study of the paedophilia domain, the experiment can be said to be hopeful in that it does not use pre-existing datasets

Table 2.1: Summary of methodologies, main findings, and limitations.

Ref.	Methodology	Main Findings	Limitations
[15]	Machine learning, Fast inference algorithm	Competitive performance, Fast inference	Slow inference with large data, Complexity of criminal network
[16]	Criminal network-based suspect prediction	Reduces execution time, Maintains competitive performance	Slow inference, Scaling challenges, Memory and time-consuming, Expensive clustering
[17]	ANOS OCL	Efficiently identifies key actors, Reveals criminal behaviors	Online networks focus, Limited offline representation, Inaccurate identification, Focus on social media
[18]	User content analysis, User network analysis, Machine learning	Effective detection in social media, Benchmark validation	Limited real-world applicability, Needs performance assessment
[19]	Suspect identification from messages, Controlled vocabulary	Promising in paedophilia	Lack of labeled data, Requires frequent updates
[20]	Feature extraction, State vector machine	Simplifies extraction, Reduces energy consumption	Challenges in traditional text classification
[21]	Centrality measures	Effective identification based on social relationships	Focuses only on social relationships, Assumes strong connections
[22]	Bi-LSTM neural network, Fuzzy c-means pre-processing	Outperforms existing methods in crime classification	Poor betweenness centrality results, Ethical implications not discussed
[23]	Controlled vocabulary	Promising in paedophilia	Lack of labeled datasets, Difficulty in supervised learning
[24]	Criminal associations, Linear association model	Effective and flexible	Ignores individual attributes, Limited comparison with one algorithm
[25]	Targeted Bayesian network	Accurate suspect identification, Anomaly detection	Limited generalizability due to specific datasets
[26]	Supervised learning, Similarity measures, K-Medoids clustering	Good results, High accuracy	Effectiveness in real-world scenarios not fully explored, Scalability and long-term performance not discussed

to identify criminal suspects. While the research considers this a disadvantage, these restrictions have been pointed out, illustrating the challenge of an accurate application of supervised machine learning in practical environments. Notwithstanding, although a controlled vocabulary method can be used to deal with the shortcomings of this approach, it still provides a viable solution to the identification of suspects on social networks. Moreover, Troncoso and Weber [24] propose a new way of uncovering associations among criminals who use social network analysis to combine individual attributes data. Their model is innovative because it brings about a new way of exposing the links within criminal organizations by highlighting its effectiveness in generating a variety of association networks. The devised model promotes the efficient momentum of investigative resources by taking the optimization of the crime planners' utility function into account. Nevertheless, the study has some drawbacks, such as the absence of an account for individual attributes, consideration of only one existing algorithm, and limited generalization since the study was conducted on only one particular dataset.

2.1. Research Gaps. Social network-based criminal suspect sensing technologies raise privacy concerns. If suitable privacy safeguards are not put in place, the collection and analysis of private data from social media networks may violate a person's right to secrecy and privacy. By ensuring the quality and consistency of the information acquired, law enforcement organizations might use this technology to access information from social media networks without endangering the privacy of people. Abuse would be prevented, the system would be easily explained and traced, and there would be sufficient numbers of controls and measures of responsibility.

Criminal Suspects Sensing is the concern of balancing between protecting people's privacy and at the same time apprehending criminals. It is possible to make certain adjustments to prevent such data collecting from being a breach of people's rights and privacy if certain gaps are closed. It is now possible to make certain that such kind of data is gathered ethically, in the right way, and in a truthful way. To address these gaps in sensing criminal suspects through social networks, it is possible to suggest the cooperation of multiple modalities. This entails establishing explicit rules and regulations, using privacy-enhancing technologies, and taking additional steps.

3. Material and Methods.

3.1. Overview of the Proposed Model. Through four main steps, this study creates a novel smart system approach for categorizing criminal suspects: ANN-based criminal suspect classification, data categorization, data encryption, and decryption phase, query-based authentication (malicious user identification), and data encryption are among the recommended approaches.

Step 1: Query-based Authentication (QBA): To ensure that only approved individuals have ownership of private data, like suspects' identities and social media profiles, QBA may be used. QBA requires users to provide specific information or answers to questions, known as queries, that only they would know. This information is compared to a pre-defined set of responses or answers associated with the user's account. The user is authenticated and granted access if the provided information matches the pre-defined answers.

Step 2: Data Categorization: User data, U^i; $i = 1, 2, \ldots, N$ is first divided into sensitive and non-sensitive categories. The data of personal and social criminal suspects are considered sensitive information S^i, as they contain information about the suspects' identity, contacts, and other private details. The instance vector q, which is a combination of personal and social data, is also considered sensitive. On the other hand, the classification result v, which indicates whether a suspect is classified as a criminal or not, is considered non-sensitive information G^i. However, its confidentiality is still important, as it should be the owner of the instance vector (police station).

Step 3: Proposed Data Encryption and Decryption Phase: The identified sensitive data S^i is encrypted using a newly introduced Two-Level hybrid ECC and RSA (ECC-RSA) Model. In the ECC-RSA model, the optimal private key $\mathrm{PK_{TC}}$ is selected via HSMEO. The encrypted data acquired from the ECC-RSA model is denoted as E^i.

Step 4: ANN-based criminal suspect classification: Criminal suspect classification can be done using ANN. A suspect's features are supplied into an ANN, classifying the suspect as either a criminal or not as an output. This integration is in consonance with the expert system approach to improve the capacity of categorizing and predicting criminal conduct.

3.2. System System. A smart system for criminal suspect sensing model has been proposed, and it is comprised of seven entities. These entities are as follows: the Key Management Centre (KMC), a Criminal Intelligence Analyst (CIA) that functions as a trusted scientific research centre, Users (suspects), a Police Station (PS), a Secure Public Server (SPS), a Social Server (SS), and a Trusted Authority (TA). The structure is divided into two different domains, namely the Preparation Domain and the Analysis Domain, each of which has a unique set of goals for the data that is being processed.

In the Preparation Domain, tasks are carried out by suspects, CIA, SS, SPS, and PS. On the other hand, in the Analysis Domain, suspects, CIA, TA, and PS (in the capacity of a query requester) are involved in complex activities. An identification method that takes into consideration both private information and social media platforms is used by the system to identify potential suspects, as shown in Figure 3.1.

KMC is a reliable organization in charge of creating keys, registering parties, and maintaining the legal parties' keys in the system. Nevertheless, the KMC is not engaged in any party or network interactions beyond this initial setup. Its only duty is to make sure the keys are managed securely and appropriately.

Users create their public and private keys and register with the Key Management Centre (KMC) in the Preparation Domain. People who are suspected or have ties to suspects are the main focus of the system. Their data is regularly sent to the Secure Public Server (SPS) via the Police Station (PS) Server. The contact information of the users—designated as c_j and c_i—is captured by their cellphones and is periodically backed up to the Social Server (SS). This information includes identification, duration, and social relationships.

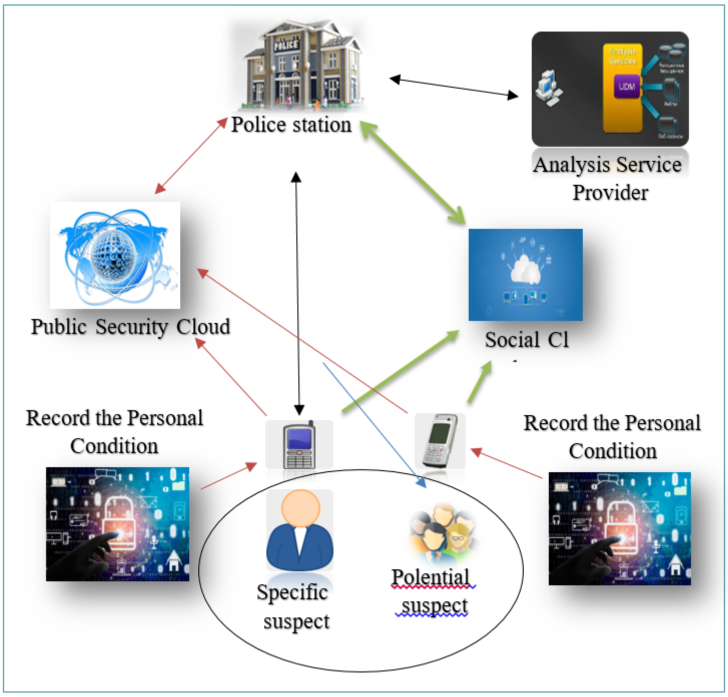
Police station
Analysis Service Provider
Public Security Cloud
Social Cl
Record the Personal Condition
Record the Personal Condition
Specific suspect
Polential suspect

Fig. 3.1: Social Network and Personal Data – based Criminal Suspect Detection System

The entities and domains of knowledge used in the proposed expert system are described and these are used to offer a holistic and efficient means of identifying criminal suspects, with the help of conventional and digital sources for improved efficiency and security.

The PS in the Preparation Domain has designated the Criminal Intelligence Analyst (CIA) as a reliable third party. Using the training datasets, it trained the tree model and saved it in the Trusted Authority (TA) for examination of criminal suspects. While the TA does the majority of the work in the Analysis Domain, the CIA may help with the categorization process. The Secure Private Server (SPS) is a server that has a large storage capacity and is responsible for collecting and storing encrypted personal data of users. Another similar server is known as the SS and it is responsible for the acquisition and archiving of encrypted social information. This information is only sent to the PS in the Analysis Domain and it helps in identifying the right products to offer to the customers. The TA in the Analysis Domain is a semi-honest participant, which has significant computing and storage resources these capabilities, process the obtained query data from the PS through operations that, if done manually, would take a lot of time. About the current system, the PS is not the only consumer of the expert system. The use of this tool is to establish if the contacts that user cj has been interacting with are suspects or not. If the decision is made that cj is a gang-affiliated criminal suspect, the PS will then pull the social details of cj from the SS. Also, the TA and CIA will investigate the criminal suspect by inputting the personal data of cj contacts and providing the encrypted results to the PS. After decrypting the result, the PS examines to identify if the contacts listed in the cj file are criminal suspects or not. The PS then communicates the result of the analysis to the police officer who has a duty of handling criminal suspects.

The most common security model in cryptography is the semi-honest model, which means that all the participants in a communication protocol adhere to the rules of this protocol but may attempt to obtain additional information. This approach focuses on the threats that originate from outside the organisation and threats that originate from within the organisation. An external attacker is an individual who attempts to penetrate the security of the data or systems to cause harm. Semi-honest is defined herein as any external invader that is untrusted and attempts to gain unauthorized access to the communication channel. This kind of attacker is sometimes referred to as an ”external eavesdropper”. Such an attacker could be able to intercept certain communication requests or answers over the shared channel and use them to get private data. Conversely, an internal attacker is a semi-honest, passive entity that has authorization to enter the network beforehand. While it will apply the expert system model, there's a possibility that it may use the information obtained from the interactions to deduce confidential information from SPS, SS, PS, or TA, among other system entities. This kind of attacker poses a serious threat to the confidentiality and privacy of the data contained in the smart system; hence adequate security measures must be implemented to prevent them. For example, access restrictions, monitoring tools, and audits may be used to prevent illegal access and identify any questionable behavior inside the system.

3.3. Query-based authentication. Query-based authentication is the first stage of the proposed smart system that attempts to verify the legitimacy of the query and the query issuer. TA is presented with a random challenge from PS at this phase. After that, TA encrypts the challenge using PS's public key and returns it to PS. PS uses its private key to decode the encrypted challenge before sending it to TA. If the decrypted assignment matches the original challenge, authentication may continue via the sorting process. Before sending the instance vector q to the TA for classification, the PS encrypts it using the public key of the TA. TA uses the encrypted model M to carry out the classification after decrypting the encrypted vector. After then, PS receives the encrypted classification outcome v and uses its private key to decode it to get the plaintext result. Social data from SS and personal data from SPS are encrypted using unique keys that are only known to the relevant organizations to guarantee security and privacy. Similarly, to ensure that only authorized personnel access the data, the access pattern is also encrypted. Moreover, the instance vector q is not reachable by TA or CIA; only PS is allowed to access it. The recommended approach applies homomorphic encryption on the encrypted instance vector q for suitable classification and outcome anonymization. This approach affirms that the classification result is encrypted and only PS has access to the result hence enhancing the security of the classification result. Further, the privacy-preserving technique retains the certainty of the classification result because homomorphic encryption preserves the accuracy of the calculation. The first type is identification using a question.

Figure 3.2 shows an example of a secure query-based authentication between the PS and the TA. It begins when PS has to come up with a challenge, encrypt it with the public key of TA, and then send it to TA. To confirm that the challenge has been changed in any way, TA then decrypts the challenge with the assistance of its private key. After this, the TA encrypts the result of the authentication process with the help of the PS's public key and sends the encrypted message back to the PS. PS then uses the private key to decrypt the encrypted authentication result.

As a result, before passing the instance vector q to TA for classification, PS encrypts q with the help of TA's public key. TA also decrypts the instance vector $\mathcal{C}$ with its private key and subsequently applies the encrypted model M on the instance, and lastly encrypts the class label Y using the PS's public key. Finally, PS employs the private key to decrypt the classification outcome to get the plaintext result.

This secure process means that all the communications and the data passed through the channels are encrypted and decrypted using the right keys to ensure that the data passed through the channel is not intercepted by unauthorized individuals and that the data passed through the channel is not altered in any way during the authentication and classification processes.

The question requester is made up of three people: It also includes TA, PS, and CIA. In this case, PS could explain to the TA that they should verify some data that contain errors. The format of this question depends on the material being taught and the structure of the course. This investigation will include any information the PS may have on the suspect including the suspect's name and ID number and any other relevant information.

Furthermore, the identification authentication is done by QBA before accessing a suspect's data. This may point to the fact that there is a need to limit the type of persons who have access to private criminal data. Besides, it could be useful in warding off negative attacks and getting unauthorized access to the system. It is also possible to categorize criminal suspects using the QBA and the same tool implies that only the law enforcement bodies can use it. This is because while the system is being used, it can prompt security questions or may request the police personnel to input more identity details. This may help to guarantee that data stored in the system is protected from other people getting access to it.

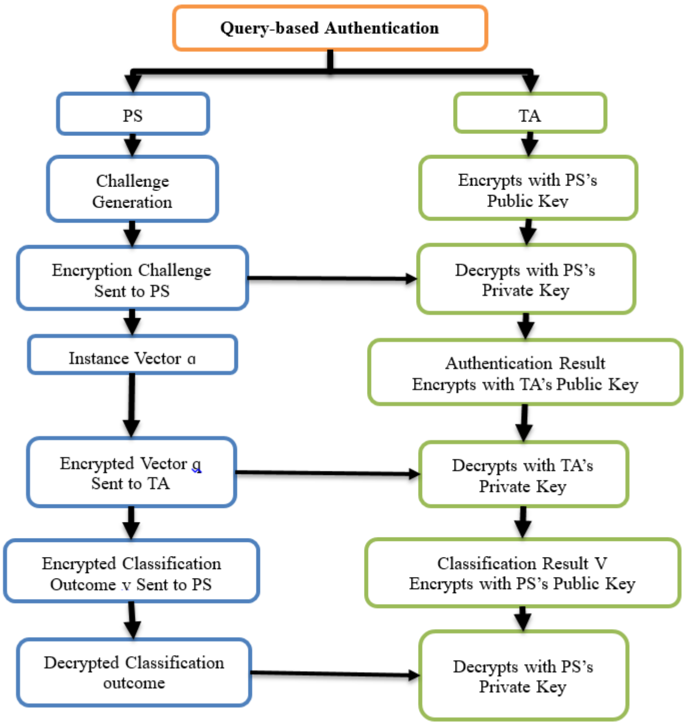
Query-based Authentication
PS
TA
Challenge Generation
Encrypts with PS's Public Key
Encryption Challenge Sent to PS
Decrypts with PS's Private Key
Instance Vector α
Authentication Result Encrypts with TA's Public Key
Encrypted Vector q Sent to TA
Decrypts with TA's Private Key
Encrypted Classification Outcome v Sent to PS
Classification Result V Encrypts with PS's Public Key
Decrypted Classification outcome
Decrypts with PS's Private Key

Fig. 3.2: Secure Query-based Authentication Process

The Basics of Using QBA to Mark Someone Who May Have Committed a Crime:

- *User authentication:* This requires the user to enter a username and password so that the user can operate the system and sort out criminal suspects.
- *Query submission:* Other details about a suspect may be required by the users if the system has authenticated the users.
- *Authentication:* Thus, the system confirms the identity of a user and allows for the assessment of information regarding a suspect. This may be achieved through biometric identification, identification by questions, and multi-level identification amongst others.
- *Suspect data access:* If the QBA is successful, then the system will allow the user access to information related to the suspect. Therefore, by applying the criminal suspect categorization system, QBA may provide a rapid and safe identification that will prevent all unauthorized individuals from accessing the details.
- *Procedure for Registration:* Customers subscribe to the data centre and input their information in the database. The client initially creates the user ID ($\mathrm{US}^{\mathrm{id}}$), then enters the password ($\mathrm{PW}^{\mathrm{id}}$) and all the user details.
- *Procedure for Login:* Following a successful registration, the user may download or upload data to the server. Without registration, no client may access the server or get data. We can prevent data loss using this strategy. Customers must first provide their login information, which consists of their password ($\mathrm{PW}^{\mathrm{id}}$) and user ID ($\mathrm{US}^{\mathrm{id}}$). After receiving it, the server checks to see whether the user is permitted. User data may be classified once the query request has been examined.

3.4. Data Categorization. Sort the user data according to the query's authenticity. The user transmits his data U^i to the network after registering with it. This information is divided into:

- Sensitive information (S^i)
- General information (G^i)

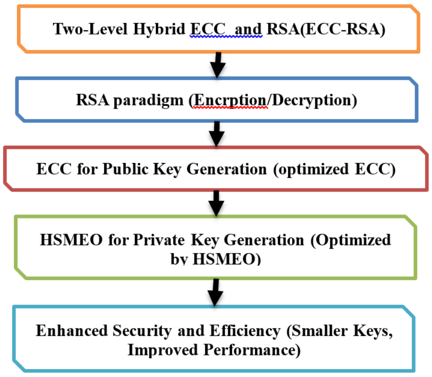
Two-Level Hybrid ECC and RSA(ECC-RSA)
RSA paradigm (Encrption/Decryption)
ECC for Public Key Generation (optimized ECC)
HSMEO for Private Key Generation (Optimized by HSMEO)
Enhanced Security and Efficiency (Smaller Keys, Improved Performance)

Fig. 3.3: Two-Level Hybrid ECC and RSA (ECC-RSA) Model

Social and personal data are two types of sensitive data. Information that may be used to identify an individual, such as name, address, phone number, and email, is referred to as personal data. This information is considered sensitive and must be protected from unauthorized access as it might be used to jeopardize the person's security and privacy. But social data also include information on the individual's social network, such as the individuals they engage with, how often and for how long, and other social connections. Since this information may reveal private facts about a person's behavior, hobbies, and connections, it may also be regarded as sensitive. Information that is not regarded as confidential and doesn't need extra security is referred to as non-sensitive information. It may include data that is freely available to the public and does not pose a threat to the security or privacy of an individual, for example, employment history information and demographic data. Given the input of sensitive and nonsensitive data, it is possible to predict the probability of a particular person being a criminal suspect. The private information is encrypted and decrypted and the data is taken as input. The parties performing the encryption and decryption job are PS, CIA, and TA.

3.5. Proposed Data Encryption/Data Decryption Model. The ECC-RSA model is a privacy-preserving method that relies on homomorphic encryption for the computation of data over encrypted data. It allows the PS to communicate with the TA and CIA through encrypted queries, thereby providing them with the encrypted data without ever exposing the plaintext to them. In this case, the encrypted output is passed to the PS to decrypt it and decide on the classification of the information without compromising the security of the information.

3.5.1. Hybrid ECC and RSA (ECC-RSA) Model. Figure 3.3 presents the Two-Level Hybrid ECC and RSA (ECC-RSA) Model which combines the RSA and Elliptic Curve Cryptosystems (ECC) for enhanced and improved cryptographic services. RSA is used for encryption and decryption since it has an excellent security strength whereas, the optimized ECC is used to generate the public key because of its efficiency and relatively small key size. The private key is then generated efficiently with the help of the new Hybrid Slime Mould Equilibrium Optimization Algorithm (HSMEO) that enhances the effectiveness of the suggested model. This makes the hybrid approach have a smaller key size and better performance compared to the other traditional schemes, hence suitable to be used in different cryptographic applications that require security and performance.

ECC and RSA are combined in a recently developed Two-Level Hybrid ECC and RSA (ECC-RSA) Model. In ECC-RSA, the RSA model is employed in the encryption/decryption of messages while the optimized elliptic curve cryptography is employed in the generation of the public key. When generating the private key, the ECC-RSA approach is again the most efficient, especially when using the new HSMEO. The cryptography methods used here are RSA and ECC-RSA though the former is slower than the later because of the larger key size. As for this, the ECC-RSA technique is more secure when it comes to smaller keys. Another factor that can help to decrease the size of the key and to create a system of secure keys is the small size of the keys of ECC. It is beneficial to combine RSA and ECC because doing so can enhance data security since it offers even smaller keys and higher efficiency. After identifying the key size, the ciphertext is applied in the encryption and decryption processes. In case ECC is employed to generate fewer but more reliable keys, the RSA encryption could require less storage space and still enhance data security. The steps of the ECC-RSA model are stated as follows: The steps of the ECC-RSA model are stated as follows:

1. Public Key Generation Using ECC. Public key cryptography is one of the most widely used methods for secure communication protocols. The discrete logarithm issue on an elliptic curve is very difficult to solve, which is one of the reasons ECC works so well. Two cryptographic keys—a public key and a private key—are created using ECC. Scale multiplying the random integer private key with a base point on this curve yields the public key. The value of the private key determines how many times the base point is multiplied by itself during the scalar multiplication operation. First, multiply a shared secret by the recipient's public key and encrypt the message using this ECC. First, the sender encrypts the message using a symmetric encryption technique like AES and the shared secret. With their private key, only the other person who is also upholding the shared secret and symmetric encryption technique may decipher the encrypted communication. Due to its small key size and excellent performance in constrained situations, ECC is often used in devices such as mobile phones and Internet of Things (IoT) devices. However, the security of ECC is affected by the parameters of the elliptic curve as well as the employment of cryptographic methods.

Step 1: Pick any prime value, v.

Step 2: To generate the public key, select a random number as $v(b)$.

Step 3: Wherein, $v(b) < v$, calculate C for the point on the curve.

Step 4: Where $C > v$, calculate the public key using Eq. (3.1).

$$p = v(b) \times C \tag{3.1}$$

where p points to the Public Key, and $v(b)$ represents the Private Key.

Step 5: The private key of ECC is generated by two different models, Lorenz Chaotic System (LCS) and SHA3-512. The generated random sequence using LCS is hashed using the SHA3-512 algorithm, expressed as per Eq. (3.2). The encryption system based on the novel key generation process in ECC achieves both speed and security by increasing the entropy of data information and reducing processing time.

$$v(b) = \text{LCS} \times \text{SHA3-512} \tag{3.2}$$

where LCS is given by:

$$\text{LCS} = \dot{x}_1 + \dot{y}_1 + \dot{z}_1 \tag{3.3}$$

and SHA3-512 indicates the hash function of SHA3-512.

Three phases, including x_1, y_1, and z_1, define the state variables and part of the LCS:

$$\begin{aligned} \dot{x}_1 &= -\sigma x_1 + \sigma y_1 \\ \dot{y}_1 &= r x_1 - y_1 - x_1 z_1 \\ \dot{z}_1 &= x_1 y_1 - b z_1 \end{aligned} \tag{3.4}$$

where σ, r, and b are real positive parameters. Depending on parameter values, the LCS can exhibit quite complex dynamics.

Step 6: Hence, the best private key is chosen from the entire generated keys using the Hybrid Slime Mould Equilibrium Optimization (HSMEO).

Step 7: Be able to output the best public key after the computations.

2. Encryption/ Decryption Using RSA. One of the most widely used public key cryptography methods, RSA is employed for protecting the content of the messages exchanged over the Internet. Two keys are generated, a public key and a private key with the help of the Random Number Generator (RSA) algorithm. These keys are created by using prime numbers and a mathematical function called modulo arithmetic. RSA is a public key cryptography algorithm that has uses in numerous fields including security of data in communication, data signatures and virtually any other use and is also used in encrypting and decrypting of data. The public key is used for the encoding of the messages while the private key is used for the decoding of the messages.

Step 1: Let us take a look at the input file.

Step 2: Now you have to add the second key, which is the public key generated through the use of the RSA algorithm, to the key pair.

Step 3: Subtract the value obtained after the RSA algorithm has been computed with the optimal value.

Step 4: This input file is then encrypted with RSA and uploaded to the server.

Step 5: After a file is uploaded, it is then downloaded from the server and decrypted from the RSA public key to get the original file.

Step 6: On the receiving end, the encrypted data is decrypted using the same public key, 'p,' to obtain the initial data.

It is proposed to employ a metaheuristic algorithm to create an optimal private key of criminals' profiles within an OSN.

3.5.2. Hybrid Slime Mould Equilibrium Optimization Algorithm (HSMEO). The HSMEO method yields the optimal private key. It is an optimizing method for challenging optimization issues that draws inspiration from nature. The motion of slime mould, a single-celled creature that looks for food sources to live, is modeled by this algorithm. Better private keys are produced by HSMEO for secure communication, ensuring the secrecy and integrity of any correspondence between two people. The other method is the HSMEO technique which is a combination of SMA and EOA. According to the HSMEO scheme, SMA is employed for search space analysis to obtain the global optimum, while EOA is employed to exploit the better settings within the region. Optimizer of equilibrium: This metaheuristic optimization tool applies the physics principle of equilibrium. The algorithm can identify the optimum response by achieving a balance between the exploration and exploitation of the search region. We employ numerous local search algorithms to identify possible geographic areas; an unordered set of moves across the solution space is investigated, and a set of potential solutions is stored in the space. This way it adapts its exploration and exploitation strategies over time based on the properties of the search space and expected solution behaviours. The parts that follow give a more detailed account of what is meant by EO's fundamental idea.

The following figure 3.4 illustrates the progressive steps of the Hybrid Slime Mould Equilibrium Optimization Algorithm (HSMEO). It starts with the general HSMEO Algorithm Phases, which describe the process of moving between different optimization phases. In Private Key Optimization, the main goal, focus is on the fact that the algorithm is based on natural processes to increase the level of protection in the process of communication. The subsequent stages describe the complex mechanisms in detail, such as the Slime Mold Motion Model, the interaction of SMA and EOA that provide a holistic optimization process, and the Equilibrium Optimizer which maintains the balance between exploration and exploitation. Local Search Algorithms enable a comprehensive search of the solution space while Continuous Strategy Modification means that the algorithm can adjust to the changes in the search space. Altogether, these phases represent the effectiveness and the high level of the HSMEO approach to address optimization problems.

Inspiration. EO starts the optimization process using the initial population. Equation (3.5) predicts that the initial quantities in the search space are produced in the following ways depending on the number of particles and the dimensions:

$$D_j^{\text{initial}} = D_{\min} + \text{rand}_j \times (D_{\max} - D_{\min}), \quad j = 1, 2, \ldots, n \tag{3.5}$$

where rand_j is a random vector in the range $[0, 1]$, D_j^{initial} defines the initial concentration vector of the i-th particle, $D_{\min}$ and $D_{\max}$ define the minimum and maximum values for the dimensions, and n is the total number of particles in the population. To classify the candidates for equilibrium, particles are sorted after being evaluated for their fitness function.

Equilibrium Pool and Candidates. For an algorithm to be more exploratory and to avoid being trapped in local optima, an equilibrium state has to be maintained during the search process. The best solutions available must be combined with new research into previously unexplored parts of the subject to maximize the effectiveness of the solutions. This strategy keeps the algorithm from becoming stuck on a less-than-ideal answer

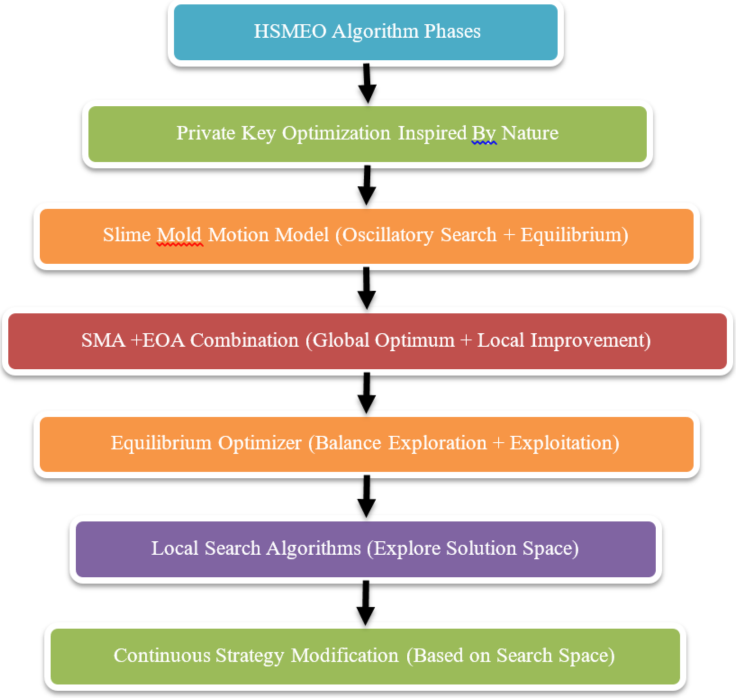
HSMEO Algorithm Phases
Private Key Optimization Inspired By Nature
Slime Mold Motion Model (Oscillatory Search + Equilibrium)
SMA +EOA Combination (Global Optimum + Local Improvement)
Equilibrium Optimizer (Balance Exploration + Exploitation)
Local Search Algorithms (Explore Solution Space)
Continuous Strategy Modification (Based on Search Space)

Fig. 3.4: Phases of the Hybrid Slime Mould Equilibrium Optimization Algorithm (HSMEO)

and allows it to keep looking for better ones. We identify five possible options in the context of equilibrium pool particles. Four of these contenders represent the best particles discovered throughout the whole optimization procedure. A particle representing the arithmetic mean of the preceding four particles is the fifth contender. While the four ideal particles aid in thoroughly exploring the search space, this fifth candidate particle aids in the exploitation of the solutions. The Hybrid Slime Mould Equilibrium Optimization (HSMEO) algorithm may provide optimum private key creation by combining these techniques.

The equilibrium pool's vector $\vec{D}_{\text{eq,pool}}$ is shown in Equation (3.6).

$$\vec{D}_{\text{eq,pool}} = \left\{\vec{D}_{\text{eq}(1)}, \vec{D}_{\text{eq}(2)}, \vec{D}_{\text{eq}(3)}, \vec{D}_{\text{eq}(4)}, \vec{D}_{\text{eq,ave}}\right\} \tag{3.6}$$

Each particle updates its concentration during an iteration by selecting random candidates among those selected with a similar probability.

Exponential Term. The EO algorithm's exploration and exploitation are balanced in part by the exponential term $\vec{F}$. According to Equation (3.7),

$$\vec{F} = e^{-\vec{\lambda}\times(T-T_0)} \tag{3.7}$$

Here, T is an iterative function that gets smaller as the number of iterations increases, and $\vec{\lambda}$ is a random vector between $[0, 1]$. As per Equation (3.8),

$$T = \left(1 - \frac{\text{Iter}}{\text{Max\_iter}}\right)^{\frac{b_2\times\text{Iter}}{\text{Max\_iter}}} \tag{3.8}$$

where Iter and Max_iter are the current and maximum number of iterations. In Equation (3.9), T_0 is calculated as,

$$\vec{T}_0 = \frac{1}{\vec{\lambda}} \ln\left[-\left(b_1 \times \text{sign}(\vec{r} - 0.5) \times \left[1 - e^{-\vec{\lambda}\times T}\right]\right) + T\right] \tag{3.9}$$

Here, the constants b_1 and b_2 regulate the capacities for exploration and exploitation, respectively. The exploration capability is stronger, and the exploitation ability is weaker with a higher value of b_1. The exploitation capability is stronger, and the exploration capability is weaker with a higher value of b_2. b_1 and b_2 are equal to 2 and 1, respectively. The direction of exploration and exploitation is indicated by $\operatorname{sign}(\vec{r} - 0.5)$. Equation (3.10) organizes and shows the final form of $\vec{F}$:

$$\vec{F} = b_1 \times \operatorname{sign}(\vec{r} - 0.5) \times \left[e^{-\vec{\lambda} \times T} - 1\right] \tag{3.10}$$

Generation Rate. By enhancing the exploitation phase, the generation rate $\vec{g}$ enables the EO algorithm to deliver precise solutions. The first-order exponential decay process is used to define the generation rate as shown in Equation (3.11),

$$\vec{g} = \vec{g}_0 e^{-\vec{\lambda} \times (T - T_0)} \tag{3.11}$$

where $\vec{g}_0$ denotes the starting point and k denotes the decay constant ($k = \lambda$). As a result, the generation rate's final expression is shown in Equation (3.12),

$$\vec{g} = \vec{g}_0 e^{-\vec{\lambda} \times (T - T_0)} = \vec{g}_0 \times \vec{F} \tag{3.12}$$

where

$$\vec{g}_0 = (\mathrm{gcp}) \times (\vec{D}_{\mathrm{eq}} - \lambda \times \vec{D}) \tag{3.13}$$

$$\mathrm{gcp} = \begin{cases} 0.5 r_1, & r_2 \geq \mathrm{gp} \\ 0, & r_2 < \mathrm{gp} \end{cases} \tag{3.14}$$

where r_1 and r_2 are random numbers between $[0, 1]$, and gcp is also known as the generation rate control parameter. The likelihood of this contribution determines the number of particles that will use generation terms to update their states. GCP is found in Equation (3.14), and its purpose is to strike a balance between exploration and exploitation with GP (GP = 0.5). A population-based optimization technique called the equilibrium optimizer works by repeatedly updating the entire population of solutions. This may result in slow convergence, particularly for complex problems or when the algorithm becomes stuck at a local optimum. Equilibrium optimizer's drawbacks include uneven exploration and exploitation, poor exploration capability, and a tendency to easily enter local optima. The population's diversity can be increased by switching from the slime mould algorithm (SMA) default simple random search approach to an equilibrium optimizer strategy.

The proposed HSMEO updating rule is shown in Equation (3.15).

$$\vec{Y}(T+1) = \begin{cases} \vec{Y}_{\mathrm{eq}}(T) + \left(\vec{Y}(T) - \vec{Y}_{\mathrm{eq}}(T)\right) \times \vec{F} + \left(\vec{g} \times \frac{1 - \vec{F}}{\lambda} \times U\right), & \mathrm{rand} < z \\ \vec{Y}_{\mathrm{eq},1}(T) + \mathrm{Levy} \times \left(\vec{X} \times \left(\vec{Y}_B(T) - \vec{Y}_A(T)\right)\right), & r < q \\ \vec{Y}(T) + \mathrm{Levy} \times \vec{Y}(T), & r \geq q \end{cases} \tag{3.15}$$

To increase local exploration, when the individuals found less information, empirical value defined as $z = 0.3$, better performance was achieved by using Levy flights instead of uniformly distributed random variables $(vb) \times \vec{vb}$, U is considered as a unit, $(\vec{Y}_{\mathrm{eq}})$ stands for a randomly chosen solution

3.6. ANN-based criminal suspect classification. Artificial neural networks are used by TA to categorize criminals in online networks. An ANN is a particular kind of machine learning algorithm that learns to predict or make judgments. The process of determining a person's probability of becoming a criminal suspect based on their age, gender, job, criminal history, and other characteristics is known as criminal suspect categorization. ANNs are very accurate in predicting if a person is likely to be a criminal suspect based on past data. A neural network is composed of layers, with neurons at the base of each layer. The categorization of a layer

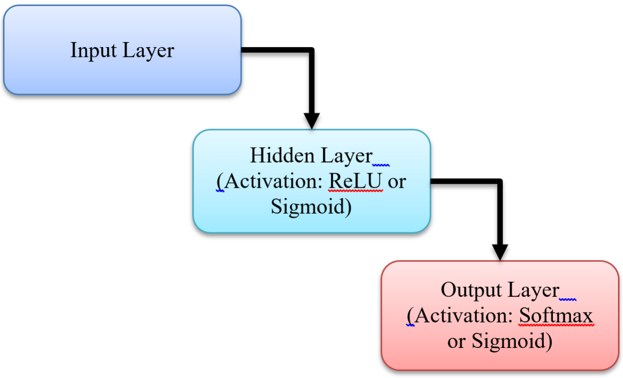
Input Layer
Hidden Layer (Activation: ReLU or Sigmoid)
Output Layer (Activation: Softmax or Sigmoid)

Fig. 3.5: Structure of a Neural Network Featuring a Hidden Layer.

as input, hidden, or output relies on its location inside the neural network. In a network, the input layer is the top layer and the output layer is the bottom layer. Hidden layers are those that are layered underneath the input and output layers. Depending on how the network is constructed, the number of neurons in each layer might change. The neurons in each layer are linked to the neurons in the layers above and below it. Complex changes take place in the network's hidden layers so that it can recognize patterns in the incoming data and generate accurate predictions.

The link between D2D and cellular channel gains lies at the core of the optimization challenge in mobile networks. It is advised to use an artificial neural network (ANN) to learn this connection and adjust the transmission intensity of D2D pairs appropriately, since this relationship cannot be deduced analytically. Stated differently, the ANN functions as a regulator, adjusting the transmission strength of D2D pairings based on data supplied to base stations about cellular channel gain. N D2D pairs are given a binary classification challenge: the optimization aims to adjust the transmission power of each pair to either $k_m = k_{\min}$ or $k_m = k_{\max}$. A fully connected ANN may be used to map cellular channel gains to the ideal binary transmission power setting for each m-th D2D pair, thereby overcoming this difficulty and optimizing the overall capacity of D2D pairings.

In Fig. 3.5, the proposed ANN for binary classification is shown. The ANN consists of an input layer $C = (C_1, C_2, \dots, C_m)$, hidden layers $S = (S_1, S_2, \dots, S_m)$, and an output layer $B = (B_1, B_2, \dots, B_m)$. The input layer of the proposed ANN aligns the cellular channel gains from the D2D users to the base stations. The input layer has an input vector of length $X \times W$ and is denoted as $\text{Out}_{c0} = (K_{1,1}, K_{1,2}, \dots, K_{X,W})$, which represents the cellular channel gains between base stations and D2D users.

The sigmoid function produces output values between 0 and 1, representing the probability that the transmission power for the m-th D2D pair should be set to 1. The transmission power for the m-th D2D pair is determined based on the output value of the ANN for that pair, as per Eq. (3.16):

$$k_m = \begin{cases} k_{\max}, & \text{if } \text{out}_B > 0.5 \\ k_{\min}, & \text{otherwise} \end{cases} \tag{3.16}$$

An activation function Φ is used to describe a basic model of a neuron by applying it to a linear combination of input data $C = (C_1, C_2, \dots, C_m)$, weights $D = (D_1, D_2, \dots, D_w)$, and a bias constant y. Consequently, the output B is such that $B = \Phi(U)$, where $U = y + \sum_{r=1}^{w} D_r C_r$. The process continues until the output layer generates the desired output by passing this output along to the subsequent layer as input data along with all of its other outputs.

The activation function Φ of the most general neuron model, the sigmoid neuron, is the sigmoid function $\Phi(U) = \frac{1}{1+e^{-\eta U}}$. The output is computed as per Eq. (3.17):

$$B = \frac{1}{1 + e^{-\eta\left(\sum_{r=1}^{w} D_r C_r - y\right)}} \tag{3.17}$$

Table 3.1: Hyperparameters Used for the Training of ANN

Parameter	**Value**
Input Layer	10
Hidden Layer	1
Hidden Neurons	10
Transfer Function	tansig
Backpropagation Network Training Function	traingdx
Epochs	100
Performance Function	MSE
Training Function	Levenberg-Marquardt Backpropagation
Output Layer	1

For a deep learning model to perform at its highest level, it is necessary to choose the appropriate hyperparameters. Some examples of the hyperparameters that influence the architecture of the model are the number of hidden layers, the number of neurons in each layer, and the learning rate. These are just a few examples. The optimal combination of hyperparameters may be discovered via the use of either a grid search or a randomized search. On the other hand, the randomized search technique chooses hyperparameters at random from a certain distribution, whilst the grid search strategy evaluates the efficacy of the model by utilizing all of the potential combinations of hyperparameters that fall within a defined range. The above result indicates that the performance of the artificial neural network (ANN) can be improved by adjusting the hyperparameters, which would improve the rate of convergence and accuracy of the model.

Table 3.1 summarizes the hyperparameters of the Levenberg-Marquardt Backpropagation-adapted hidden layer with ten neurons. Ten computational elements form the intermediate layer that uses the tansig transfer function to transform the input to an output value between -1 and 1. Regarding backpropagation training, there is only one function, and it is called Traingdx. It modifies the net weights to minimize the gap between the outcome predicted and the actual outcome by using the back propagational technique with the inclusion of the momentum factor. The output layer has one node for ten input variables and the mean squared error is used to check the actual and predicted values after 100 epochs of training. Last but not least, it would be possible to state that ANNs can contribute to the improvement of the accuracy and efficiency of the categorization of criminal suspects. The effectiveness depends on the quality and relevance of the training data, as well as the ability to analyze the data. It is quite possible that integrating human and legal approaches with ANN may help criminal justice systems to better and more logically sort suspects. The research findings will be provided to a law enforcement agency if the user does not meet the criteria of being suspicious. It will also be possible for the agency to look for such data and get the list of possible suspects who have not been arrested yet.

4. Experimental Results. The study used query-based authentication (QBA) to ensure that only authorized users with proper identification details accessed relevant information. Machine intelligence was applied to categorize data into two groups: personal and social criminal data which can be considered as sensitive data and classification results which can be considered as non-sensitive data. This paper also presented the Two-Level hybrid ECC and RSA (ECC-RSA) encryption model and compared it with other traditional methods such as AES, DES, RSA, and ECC and proved that the proposed model was more secure and efficient. The encryption/decryption time and throughput analysis of ECC-RSA were quite satisfactory. In this study, the HSMEO algorithm improved the ECC-RSA model, and it was found that the key generation time of the proposed algorithm was smaller than SMO, EO, PSO, MFO, and FFO. Graphical illustrations corroborated the viability of the ECC-RSA encryption model as a more secure and optimally performing encryption technique compared to the others. Finally, the findings of the experiment show that the smart system methodology above is effective in concealing the data from other users and in categorizing criminal suspects using an expert system approach.

4.1. Experimental Setup. This recommended solution has been implemented in the MATLAB environment. The Two-Level Hybrid ECC and RSA (ECC-RSA) Model utilizes a classifier-based artificial neural network and a hybrid slime mould optimization for suspect profiling. In this part, the model analysis is presented more graphically as can be seen below. The following are some of the existing approaches that can be compared to the proposed methodology: Among these metaheuristic algorithms, Slime Mould Optimisation (SMO) [26], Equilibrium Optimizer (EO) [27], Particle Swarm Optimisation (PSO) [28], Moth Flame Optimizer (MFO) [29], and Firefly optimizer FFO [30]. They are also compared with RSA [32], Elliptic Curve Cryptography [31], AES [33], and Information Encryption Standard (DES) [34]. That is why measures such as the time taken to generate keys, encrypt, and decrypt are used to assess the recommended approach.

4.2. Performance Metrics. A performance matrix is a set of parameters of a measurable nature that have been formulated to measure the performance, effectiveness, and quality of an encryption system. These are generation time, encryption time, decryption time, security strength, trust score, throughput, and delivery ratio. These metrics include aspects of the encryption system in terms of speed, reliability, security, and efficiency in processing and transferring encrypted data. The performance matrix is a means of comparing the performance of the system in the encryption of data to make the right decision about whether the system is efficient in protecting sensitive information or not.

4.2.1. Key Generation Time (Sec). Key generation time is the time taken to generate the keys that are used in the encryption and decryption of information. It is an important parameter as shorter key generation time reduces the time that is taken to set up secure communication and protect data to start secure transactions and communication.

The formula to estimate the KGT is described in Equation (4.1).

$$\mathrm{KGT} = f(P, A, B, G, n, h) \tag{4.1}$$

4.2.2. Encryption Time (sec). Encryption time is the time required for converting the plain text data into cypher text with the help of the encryption algorithm. This metric is very vital, especially in areas where there is a need to minimize the time taken in the transfer of data. Quicker encryption times can be used in the protection of data in a very efficient manner without much or any impact on the system. The formula to calculate the encryption time is as follows Equation 4.2.

The formula to calculate the encryption time is as follows (Equation (4.2)):

$$\mathrm{ET} = \left(\frac{\mathrm{psize}}{\mathrm{ksize}}\right)^2 \cdot \mathrm{KGT} + \mathrm{et} \cdot \mathrm{psize} \tag{4.2}$$

where ET is the time required to encrypt one byte of the plaintext message using RSA encryption in seconds; KGT is the size of the RSA key in bits; and psize is the size of the plaintext message to be encrypted in bytes.

4.2.3. Decryption Time (sec). The decryption time is the amount of time taken to decrypt the given ciphertext to plaintext using a decryption algorithm. This metric is very useful in the timely retrieval and use of encrypted information. Less decryption time implies that the secured information is processed and made available more quickly; thus, there is improvement in the processing of data and decision-making.

The formula for calculating the decryption time is represented in the following equation:

$$\mathrm{Dt} = \left(\frac{\mathrm{csize}}{\mathrm{ksize}}\right)^2 \times \mathrm{kgt} + \mathrm{dt} \times \mathrm{csize} \tag{4.3}$$

where dt is the quantity of time it takes to decode one byte of ciphertext using RSA decryption in seconds, and csize is the extent of the ciphertext message to be decrypted in bytes.

4.2.4. Security. In the context of encryption systems, security is the capacity of encryption algorithms to prevent unauthorized use of data or alteration of it. The security strength of the encryption scheme depends on the types of attacks that can be carried out on the cipher system. High security also plays a role in preventing the interception of encrypted data by individuals who are not authorized to access it, and it also protects the data from being tampered with or from loss of integrity.

It is possible to state that the level of security provided by a cryptographic approach is equal to $2^{(n-1)}$, where n is the number of bits in the key. For example, an encryption technique with a key of 128 bits would be as secure as $2^{(128-1)} = 2^{127}$, which is approximately 1.7×10^{38}. Thus, it is thought that such encryption is essentially impossible to crack with present-day techniques, even with 2^{127} different keys.

4.2.5. Trust Score. Trust score is another invented measure that tries to express the degree of trust and reliability of an encryption system. It depicts the degree of confidence that the system can protect sensitive information from any danger. A high trust score implies that the encryption system is perceived to be effective in the protection of information and that the information is secure from other individuals' interference. Equation 4.4 shows an example of a simple formula for computing the trust score. Thus, the Trust score is determined by the following equation:

$$\text{Transparency} \times 0.1 + \text{Competence} \times 0.2 + \text{Honesty} \times 0.3 + \text{Reliability} \times 0.4 \tag{4.4}$$

Each element in Eq. The weight of each factor, in this case, Equation 4.4 is determined by the overall confidence that has been set. It was seen that reliability, which can be defined as the extent to which a firm fulfils its commitment, was the most heavily weighted at 0.4 on the other hand, the level of transparency which measures the extent to which a company is willing to reveal information on the actual business processes has the least scale weight of 0.1. This is because it is possible to obtain the specific ratings of each piece by using different metrics like the effectiveness statistics, the user feedbacks and the third-party reports. These are then added to the corresponding weight to arrive at the total trust score.

4.2.6. Throughput. Throughput is defined as the ability of an encryption system to process the data or data transfer rate in a certain period. It defines the efficiency by which the system can perform data encryption and decryption tasks. This is because data can be processed and transmitted in a fast manner to ensure that it does not become a bottleneck and to ensure that the efficiency of applications and systems is not compromised. Equation 4.5 may be used to represent the throughput formula:

$$\text{Throughput} = \frac{\text{Total amount of data transmitted or processed}}{\text{Total time taken}} \tag{4.5}$$

Quite often, the answer to this equation is presented in terms of bits per second (bps), bytes per second (Bps), or another data transfer rate unit.

4.2.7. Delivery Ratio. The delivery ratio is the ratio of the number of packets of encrypted data successfully delivered to the total number of packets transmitted through the encryption system. It shows the efficiency of the system in ensuring that protected data is delivered to intended users. A high delivery ratio proves the efficiency of the system in providing secure delivery of encrypted data and guarantees effective communication and data transfer. Equation 4.6 is used to compute the packet delivery ratio.

$$\text{PDR} = \left(\frac{\text{Number of packets received at destination}}{\text{Total number of packets sent from source}}\right) \times 100\% \tag{4.6}$$

4.3. Results. Table 4.1 presents AES, DES, RSA, and ECC along with the recommended Two-Level hybrid RSA and ECC (ECC-RSA) model's security, trust score, throughput, encryption and decryption times, and delivery ratio. When compared to other methods, its security percentage of 98.53% offers the best degree of protection [35]. It is regarded as the most trustworthy algorithm overall, with the highest trust score of 4.83. The quickest encryption and decryption times for the suggested ECC-RSA approach are 3.459 and 2.994 seconds, respectively [36]. The ECC algorithm has the quickest decryption time (5.1243 seconds) and the slowest encryption time (3.7086 seconds) among the various approaches. The highest throughput of 77.57 for the suggested ECC-RSA methodology illustrates even more how much quicker it can process larger quantities of data than the other methods [37]. With the best delivery ratio of all the algorithms, 0.9884, it provides even more proof of a better chance of an efficient message transfer. The adoption of the extremely secure and efficient ECC-RSA method is recommended. The high degree of security and quick encryption and decryption operations of the proposed approach may be advantageous for classifying criminal suspects.

Table 4.1: Comparison of Overall Performance

Techniques	**Encryption Time (Sec)**	**Decryption Time (Sec)**	**Security (%)**	**Trust Score**	**Throughput**	**Delivery Ratio**
Proposed ECC-RSA	3.459	2.994	98.53	4.83	77.57	0.9884
ECC	3.7086	5.1243	92.64	4.36	69.354	0.9234
RSA	4.1348	4.793	94.54	4.11	65.45	0.9614
AES	9.6512	6.6401	85.669	3.07	39.94	0.8823
DES	6.182	11.9697	72.45	3.2	44.35	0.8534

Table 4.2: Comparison of Key Generation Time (HSMEO)

Methods	**Key Generation Time (Sec)**
Proposed HSMEO	1.97545
SMO	3.13684423
EO	2.708879857
PSO	3.891071564
MFO	4.668197068
FFO	5.41265866

Table 4.2 displays the significant production times for SMO, EO, PSO, MFO, and FFO in the developed HSMEO model. Unlike the previous models, the proposed HSMEO model produces keys in 1.97545 seconds or less. It has been shown that the generation of keys takes longer for SMO, EO, PSO, MFO, and FFO [38], with MFO and FFO taking much longer. These findings suggest that the HSMEO model generates keys faster than other methods do, according to the proposed model. The last aspect of the suggested model is that it takes more time to generate compared to other models, which is crucial when grouping criminal suspects. Because secure jobs may often be completed more quickly, the proposed model is more beneficial. The conclusions of the study are significant because applications may become safer and more effective if there are shorter key generation periods [39]. To perform the encryption and decryption functions, a key is mandatory. The improved performance of the HSMEO technique has a significant impact on the design and construction of secure cryptographic systems. Given all this, the study reveals that the HSMEO approach can be used to quickly derive safe encryption keys. In this part, the authors compare the graphical representation of the metrics with other techniques like SMO, EO, PSO, MFO, and FFO. This research also compared and contrasted various key generation efficiency strategies with some of the most commonly used encryption algorithms like RSA, DES, AES, and ECC. A comparison of these approaches was also conducted simultaneously.

In addition to the five other optimization strategies (SMO, EO, PSO, MFO, and FFO), Figure 4.1 illustrates the amount of time required to generate a key for the HSMEO strategy that is suggested. Given the statistics, the HSMEO approach that was presented is superior to other methods in terms of its effectiveness since it only takes 1.97545 seconds to produce a key. The EO, SMO, PSO, MFO, and FFO strategies, with the respective key creation times being 2.84568425 seconds, 3.13684423 seconds, 3.891071564, 4.668197068, and 5.41265866 seconds, respectively. In addition to this, it emphasizes the fact that the HSMEO technique offers superior performance in comparison to other approaches [40]. It is quite important to choose the optimization approach that is designed to be the most efficient to generate safe encryption keys. The outcomes of the investigation are supported by the graphical depiction of the comparison of the times at which the key was generated.

Figure 4.2 displays a visual depiction of the Encryption Time (in seconds) for the four widely used encryption methods (RSA, ECC, AES, and DES) as well as the proposed ECC-RSA approach. According to the graph, the quickest encryption time was 3.708 seconds for ECC, while the second-fastest encryption time was 3.459 seconds for the Proposed ECC-RSA technique. The encryption timings for RSA, DES, and AES were 4.134, 6.182, and 9.651 seconds, in that order. It is fast and simple to see how different encryption methods

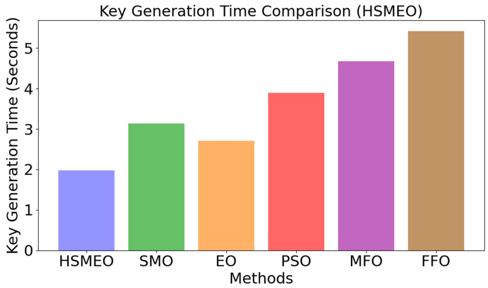

Fig. 4.1: Comparison of key generation time of the proposed method with existing techniques.

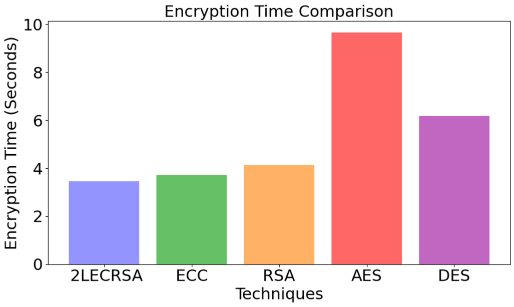

Fig. 4.2: Comparison of encryption time of the proposed method with existing techniques.

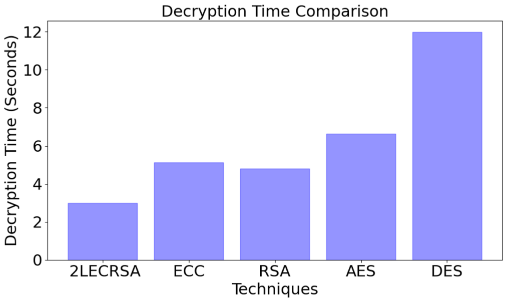

Fig. 4.3: Comparison of decryption time of the proposed method with existing techniques.

vary in terms of encryption time thanks to the graphical depiction. It emphasizes how the Proposed ECC-RSA method and ECC work comparatively well in comparison to the other encryption algorithms, especially AES, which has the slowest encryption time.

Figure 4.3 shows that, at 2.994 seconds, the Proposed ECC-RSA technique has the quickest decryption time. The next four algorithms RSA, ECC, AES, and DES have decryption times of 4.793, 5.124, 6.640, and 11.969 seconds, respectively. The suggested ECC-RSA technique highlights a large efficiency gain when compared to conventional encryption algorithms, particularly DES and AES, which have the fastest decryption times.

The comparison of the suggested ECC-RSA methodology and the four most commonly used encryption methods (RSA, AES, DES, and ECC) is depicted in Figure 4.4. The recommended ECC-RSA technique demonstrated a high level of security in preventing unauthorized access to sensitive information with the security percentage of 98%. 53%. The security percentages that ECC, RSA, and AES obtained—92.64%,

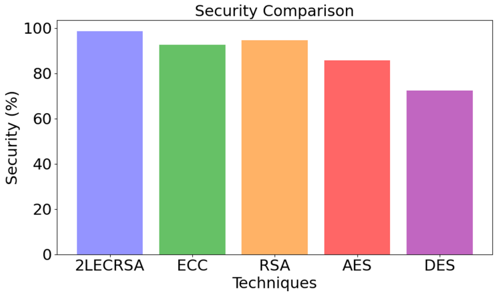

Fig. 4.4: Comparison of security of the proposed method with existing techniques.

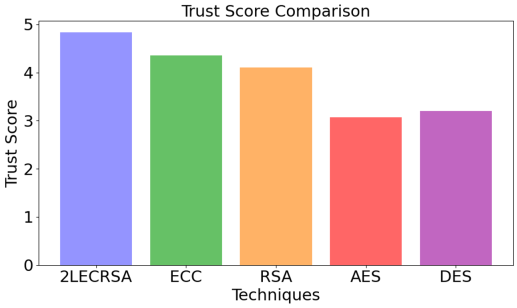

Fig. 4.5: Comparison of the trust score of the proposed method with existing techniques.

94.54%, and 85.69% also demonstrate the level of protection. On the other hand, the security percentage of the DES technique was only 72.45% which is a little less than the other methods in terms of security. As a consequence, it may be more vulnerable to invasions and unauthorized access to confidential information.

This is why the trust score is a parameter that should be taken into consideration when choosing the encryption algorithm. The trust score is an algorithm that measures the confidence that users have in the ability of the algorithm to secure their information. The ECC-RSA encryption method is proposed to be implemented as shown in figure 4.5 below as well as the graphical representation of the trust ratings of the different encryption methods. The highest trust score of 4 is achieved among all methods. Based on the foregoing analysis, the recommended ECC-RSA algorithm is a good encryption for protecting sensitive information due to its flexibility. Examples of trust-related characteristics that, depending on the specific use case, are often given different weights in trust score calculations include openness, honesty, reliability, and credibility.

A crucial metric called throughput measures how fast an encryption and decryption method can process data. A high throughput rate algorithm can process enormous volumes of data quickly and effectively. Figure 4.6 illustrates throughput graphically. With a 77.57 throughput score, the proposed ECC-RSA method proved to be capable of handling large amounts of data quickly. ECC, RSA, and AES have correspondingly quite high throughput scores of 69.35, 65.45, and 39.94. However, the DES technique came in last with a throughput score of 44.35, processing data far more slowly than the other methods. Selecting an encryption method requires careful consideration of both security and performance. Robust security is necessary to avoid issues with system performance, but it's also critical to make sure the algorithm can handle data rapidly. The security of the confidential data and the efficiency of the system can be preserved by implementing the suggested ECC-RSA approach as it exhibits high security and overall throughput performance.

The delivery ratio is a parameter that estimates the number of data packets that can be transmitted by encryption. A high delivery ratio means that the data packets are sent without errors, are dependable and are delivered often. As aforementioned, the delivery ratio was 0.9884, the ECC-RSA approach recommended herein proved to offer the highest level of dependability when it comes to the transportation of data packets

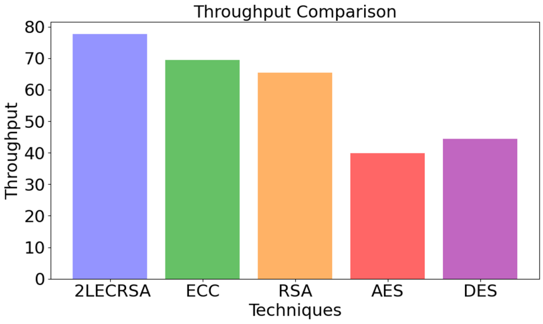

Fig. 4.6: Comparison of throughput of the proposed method with existing techniques.

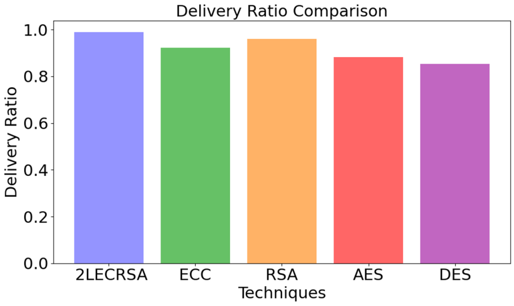

Fig. 4.7: Comparison of the Delivery Ratio of the proposed method with existing techniques.

(as shown in FIgure 4.7).In this regard, the delivery ratio is a significant parameter that should be considered in the selection of an encryption scheme because it influences the reliability and speed of the data transmission. As much as the suggested ECC-RSA technique has been proven to be effective on all the parameters including delivery ratio, it can be safe and reliable for data transmission.

5. Discussion. From the results of the study, it can be concluded that the ECC-RSA model is more secure and has a higher trust score, throughput, encryption and decryption time, and delivery ratio as compared to other encryption techniques and thus, can be used in real-time applications where data security is a major concern. The key generation method HSMEO used in the ECC-RSA model also turns out to be efficient as compared to other key generation methods as it generates keys in less time than the other methods which proves that the proposed ECC-RSA model is efficient. The use of graphs for the metrics also helps enrich the study's result since it also illustrates how the proposed ECC-RSA model is more secure, efficient, and effective compared to the other encryption methods. Therefore, it can be concluded that the ECC-RSA model described in the paper, supported by the machine intelligence technique, can be recognized as a secure, efficient, and reliable tool for ensuring the confidentiality of information in such applications as Internet banking, e-commerce, and real-time data transfer. The incorporation of Smart System and Expert System in this model also improves the model's ability to address data security issues efficiently.

5.1. Practical implications. The practical relevance of HSMEO for the identification of criminal suspects through the artificial intelligence network is connected with law and security. This method is unique and well-defined to the use of machine learning algorithms in the processing of criminal and social data from the police information system, and social media among other sources. For the encryption and decryption processes, a two-level hybrid encryption method, which is ECC-RSA, is utilized. This division of the system guarantees the isolation of confidential and non-confidential data, thus increasing security and privacy. On the other hand, the inclusion of HSMEO for private key selection is a quick unique innovation in the encryption processes, to make sure that they will be efficient and secure. In addition, the ANN classifier which can be used to differentiate criminal suspects in social networks with a much higher level of accuracy and realness is another important implementation. Practically, these findings may enrich the existing toolkits of law enforcement agencies allowing them to combat cybercrimes more efficiently and realistically. The utilization of the proposed model in MATLAB's environment signifies the practicality of the real-world applications.

5.2. Limitations. The proposed hybrid Lime Mould equilibrium optimization-based artificial neural network approach for criminal suspect identification online crime has a lot of potential in solving online crimes. Nevertheless, its deficiency is significant. On the one hand, it is largely dependent on the credibility and accessibility of crime and social statistics, a condition that would lead to inaccuracies if either limited or biased. Additionally, we have the issue of using machine learning methods which can generate bias and interpretability problems, and, eventually, classify unfairly.

6. Conclusion. This research uses machine learning methods to categorize criminal suspects on social media. The proliferation of virtual social networks has increased the frequency of communication between suspects in criminal activities. On the other hand, not enough focus has been placed on how criminal suspects are categorized on social media. In an attempt to close this gap, this study proposed a novel method of identifying criminal suspects using data from social media and police information systems. The four primary phases of the suggested approach are query-based authentication, data categorization, suggested data encryption and decryption, and criminal suspect classification based on ANN. As part of the data categorization procedure, sensitive and non-sensitive data were separated into many categories. HSMEO chose the best private key by using the ECC-RSA Model, a two-level hybrid of ECC and RSA, to encrypt and decode data. The criminal suspects on the social network were categorized by an ANN. The application of the smart system approach is based on the MATLAB platform, which enables a large number of users to participate in the system testing. Our study's unique method for categorizing criminal suspects on social media will have a significant impact on law enforcement agencies. The efficacy and efficiency of the offered approach indicate that it might be highly beneficial in preventing cybercrime. The use of the expert systems and MATLAB made it possible for the researcher to conduct an efficient and fast assessment of the performance of the system. The research contrasted the ECC-RSA system under investigation's total key generation effectiveness. Based on the results, the system's high degree of security (98.53) indicates that it is resistant to potential security breaches. The system's 77.57 throughput score demonstrated that it could easily handle a large volume of data.

This research lays the foundation for further exploration in the application of machine learning for criminal suspect classification on social media. Future work could focus on expanding the dataset to include more diverse social media platforms and integrating real-time data streams for more immediate identification of suspects.

Availability of Data and Materials:. The data that support the findings of this study are available on request from the corresponding author.

REFERENCES

[1] Golf-Papez, Maja, and Ekant Veer. ”Feeding the trolling: Understanding and mitigating online trolling behavior as an unintended consequence.” Journal of Interactive Marketing 57, no. 1 (2022): 90-114.

[2] Adisa, Omolola Tobiloba. ”The impact of cybercrime and cybersecurity on Nigeria's national security.” (2023).

[3] Ch, Rupa, Thippa Reddy Gadekallu, Mustufa Haider Abidi, and Abdulrahman Al-Ahmari. ”Computational system to classify cybercrime offenses using machine learning.” Sustainability 12, no. 10 (2020): 4087.

[4] Bhebe, Qinisani Phambili. ”Social Media and National Security in Sub-Saharan Africa: the case of Zimbabwe (2000-2017).” PhD diss., University of Johannesburg, 2022.

[5] Kobia, Robert. ”International Inter-Agency Coordination of State and Non-State Actors in Combating Global Cyber Threat: Case Study of Kenya and Zambia.” PhD diss., University of Nairobi, 2021.

[6] Camacho, David, Angel Panizo-LLedot, Gema Bello-Orgaz, Antonio Gonzalez-Pardo, and Erik Cambria. ”The four dimensions of social network analysis: An overview of research methods, applications, and software tools.” Information Fusion 63 (2020): 88-120.

[7] Nassif, Ali Bou, Manar Abu Talib, Qassim Nasir, and Fatima Mohamad Dakalbab. ”Machine learning for anomaly detection: A systematic review.” Ieee Access 9 (2021): 78658-78700.

[8] Sviatun, O. V., O. V. Goncharuk, Chernysh Roman, Olena Kuzmenko, and Ihor V. Kozych. ”Combating cybercrime: economic and legal aspects.” WSEAS Transactions on Business and Economics 18 (2021): 751-762.

[9] Hasan, Mohammad Kamrul, Zhou Weichen, Nurhizam Safie, Fatima Rayan Awad Ahmed, and Taher M. Ghazal. "A Survey on Key Agreement and Authentication Protocol for Internet of Things Application." IEEE Access (2024).

[10] Quinn, Paul, and Gianclaudio Malgieri. "The difficulty of defining sensitive data—The concept of sensitive data in the EU data protection framework." German Law Journal 22, no. 8 (2021): 1583-1612.

[11] Afshin Faramarzi, Mohammad Heidarinejad, Brent Stephens, Seyedali Mirjalili, Equilibrium optimizer: A novel optimization algorithm, Knowledge-Based Systems, Volume 191, 2020, 105190, ISSN 0950-7051,

[12] Bansal, Malti, Shubham Gupta, and Siddhant Mathur. "Comparison of ECC and RSA algorithm with DNA encoding for IoT security." In 2021 6th international conference on inventive computation technologies (ICICT), pp. 1340-1343. IEEE, 2021.

[13] Shimin Li, Huiling Chen, Mingjing Wang, Ali Asghar Heidari, Seyedali Mirjalili, Slime mould algorithm: A new method for stochastic optimization, Future Generation Computer Systems, Volume 111, 2020, Pages 300-323, ISSN 0167-739X,

[14] Ashtiani, Matin N., and Bijan Raahemi. "Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review." Ieee Access 10 (2021): 72504-72525.

[15] Jhee, Jong Ho, Myung Jun Kim, Myunggeon Park, Jeongheun Yeon, Yoonshin Kwak, and Hyunjung Shin. "Fast prediction for suspect candidates from criminal networks." In 2023 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 353-355. IEEE, 2023.

[16] Jhee, Jong Ho, Myung Jun Kim, Myeonggeon Park, Jeongheun Yeon, and Hyunjung Shin. "Fast Prediction for Criminal Suspects through Neighbor Mutual Information-Based Latent Network." International Journal of Intelligent Systems 2023 (2023).

[17] Chachoo, Manzoor Ahmad. "Social network analysis based criminal community identification model with community structures and node attributes." In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 334-339. IEEE, 2022.

[18] Gupta, Atika, Priya Matta, and Bhasker Pant. "Identification of cybercriminals in social media using machine learning." In 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), pp. 1-6. IEEE, 2022.

[19] Florentino, Érick S., Ronaldo R. Goldschmidt, and Maria Cláudia Cavalcanti. "Identifying Suspects on Social Networks: An Approach based on Non-structured and Non-labeled Data." In ICEIS (1), pp. 51-62. 2021.

[20] Shafi, Imran, Sadia Din, Zahid Hussain, Imran Ashraf, and Gyu Sang Choi. "Adaptable reduced-complexity approach based on state vector machine for identification of criminal activists on social media." IEEE Access 9 (2021): 95456-95468.

[21] Alebouyeh, Zeinab and Amir Jalaly Bidgoly. "Criminals Detection in Social Networks Using Centrality Measures Algorithm." Jordan Journal of Electrical Engineering (2021).

[22] Deepak, Gerard, S. Rooban, and A. Santhanavijayan. "A knowledge centric hybridized approach for crime classification incorporating deep bi-LSTM neural network." Multimedia Tools and Applications 80, no. 18 (2021): 28061-28085.

[23] Florentino, Érick S., Ronaldo R. Goldschmidt, and Maria Claudia Cavalcanti. "Identifying criminal suspects on social networks: A vocabulary-based method." In Proceedings of the Brazilian Symposium on Multimedia and the Web, pp. 273-276. 2020.

[24] Troncoso, Fredy, and Richard Weber. "A novel approach to detect associations in criminal networks." Decision Support Systems 128 (2020): 113159.

[25] Gruber, Aviv, and Irad Ben-Gal. "Using targeted Bayesian network learning for suspect identification in communication networks." International Journal of Information Security 17 (2018): 169-181.

[26] Bharathi, S. T., B. Indrani, and M. Amutha Prabakar. "A supervised learning approach for criminal identification using similarity measures and K-Medoids clustering." In 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp. 646-653. IEEE, 2017.

Edited by: Manish Gupta
Special issue on: Recent Advancements in Machine Intelligence and Smart Systems
Received: Aug 16, 2024
Accepted: Nov 1, 2024