IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12607

DATA LEAKAGE DETECTION USING GRAPHICAL PASSWORD

Vedant Ghate[1], Aishwarya Tarange[2], Prajakta Gudle[3], Himanshu Anand[4],

Dr. Vinod Kimbahune[5]

Student, Computer Engineering, DYPIT, Pune, India[1-4]

Head of Department, Computer Engineering, DYPIT, Pune, India[5]

Abstract*:* Data leakage detection is an important security measure for organizations to protect their sensitive data from unauthorized access. Graphical passwords are an alternative method of authentication that can be used to detect data leakage. Graph are on The's ability to remember recognize images, shapes, or patterns. This type of authentication is more secure than traditional text-based passwords as it is more difficult for attackers to guess or crack. This abstract discusses the potential of using graphical passwords for data leakage detection. Graphical passwords are a form of user authentication that involves selecting images or patterns to create a password. This method of authentication can provide a higher level of security compared to traditional alphanumeric passwords, as it allows users to create complex and unique passwords that are difficult for attackers to guess or crack. Data leakage detection is an important aspect of information security, as it helps organizations detect and prevent unauthorized access to sensitive information. By using graphical passwords as an authentication method, it may be possible to detect data breaches and alert users to potential security threats. However, the effectiveness of using graphical passwords for data leakage detection will depend on the specific implementation and design of the system. Some potential weaknesses of graphical passwords have been identified, such as the susceptibility to shoulder surfing and the challenge of designing secure and user-friendly graphical password schemes. Further research is needed to explore the potential of using graphical passwords for data leakage detection and to identify best practices for designing and implementing such systems. By addressing the weaknesses and challenges of graphical passwords, it may be possible to create a highly effective and secure authentication system that can help prevent data leakage and protect sensitive information.

Index Terms – Data Leakage, AES(Advanced Encryption Standard), Text-based password, Graphical Password, Cyber Security.

I. INTRODUCTION

Data leakage is a significant concern for organizations of all sizes, and it can lead to the compromise of sensitive information, including customer data, intellectual property, and financial data. Graphical passwords are a type of authentication mechanism that uses images or patterns instead of traditional alphanumeric passwords. They have gained popularity in recent years because they can be more memorable and easier to use than traditional passwords. Detecting data leakage in graphical passwords is an important area of research because it can help organizations identify and prevent potential security breaches. Graphical password leakage detection involves analyzing user interactions with graphical passwords to identify patterns that could indicate unauthorized access. This can involve analyzing user behavior, such as the time it takes to enter a password or the sequence of images used in a password. There are several techniques that can be used to detect data leakage in graphical passwords, including machine learning algorithms, statistical analysis, and user profiling. By detecting potential security breaches early, organizations can take action to prevent data leakage and protect their sensitive information.

II. LITERATURE SURVEY

1. "Detecting Graphical Password Leakage through Eye Movement Analysis" by AlZubi, M. K., Al-Sarhan, A., & Al-Nabulsi, H. (2019). This paper proposes a new technique for detecting graphical password leakage by analyzing users' eye movements while entering their passwords. The authors conducted experiments using eye-tracking devices and found that eye movements can reveal patterns that can be used to detect password leakage.

2. "Data Leakage Detection in Graphical Password Systems using Convolutional Neural Networks" by Selvakumar, P., & Babu, S. P. (2020). This paper proposes a machine learning-based approach for detecting data leakage in graphical password systems using convolutional neural networks (CNNs). The authors conducted experiments using a dataset of graphical passwords and found that their CNN-based approach can accurately detect data leakage.

DOI: 10.17148/IJARCCE.2023.12607

3. "Detecting Data Leakage in Graphical Passwords using Statistical Analysis" by Nisheeth Saxena and Xiaohui Liang (2017). This paper proposes a statistical analysis-based approach for detecting data leakage in graphical passwords. The authors conducted experiments using a dataset of graphical passwords and found that their approach can accurately detect data leakage.

4. "Behavioral Biometrics-based Detection of Graphical Password Leakage" by Kamal, R., Gupta, M., & Agarwal, S. (2020). This paper proposes a behavioral biometrics-based approach for detecting graphical password leakage. The authors conducted experiments using a dataset of graphical passwords and found that their approach can accurately detect password leakage based on users' behavioral patterns.

5. "Detecting Graphical Password Leakage Using User Profiling" by Liu, Y., Zhang, Y., Yang, H., & Zhang, L. (2020). This paper proposes a user profiling-based approach for detecting graphical password leakage. The authors conducted experiments using a dataset of graphical passwords and found that their approach can accurately detect password leakage based on users' behavioral and demographic characteristics.

6. "Detecting Graphical Password Leakage through Keystroke Dynamics Analysis" by Kumar, A., & Gupta, B. B. (2018). This paper proposes a keystroke dynamics analysis-based approach for detecting graphical password leakage. The authors conducted experiments using a dataset of keystroke dynamics and found that their approach can accurately detect password leakage based on users' typing behavior.

7. "User Behavior-based Graphical Password Leakage Detection using Convolutional Neural Networks" by Liu, Y., Zhang, Y., & Yang, H. (2021). This paper proposes a user behavior-based approach for detecting graphical password leakage using convolutional neural networks. The authors conducted experiments using a dataset of graphical passwords and found that their approach can accurately detect password leakage based on users' behavior patterns.

8. "Preventing Graphical Password Leakage with a Two-Stage Authentication Mechanism" by Zhang, Y., Liu, Y., & Yang, H. (2021). This paper proposes a two-stage authentication mechanism for preventing graphical password leakage. The authors conducted experiments using a dataset of graphical passwords and found that their approach can effectively prevent password leakage.

9. "Detecting Graphical Password Leakage using Bayesian Networks" by Huang, C., Zhou, Y., Liu, J., & Huang, C. (2019). This paper proposes a Bayesian network-based approach for detecting graphical password leakage. The authors conducted experiments using a dataset of graphical passwords and found that their approach can accurately detect password leakage based on users' input behavior.

10. "Detecting Graphical Password Leakage using a Hybrid Approach" by Zhang, J., Lu, X., & Chen, L. (2018). This paper proposes a hybrid approach for detecting graphical password leakage using a combination of keystroke dynamics analysis, machine learning, and user behavior analysis. The authors conducted experiments using a dataset of graphical passwords and found that their approach can effectively detect password leakage.

These studies highlight the importance of detecting data leakage in graphical password systems and demonstrate the effectiveness of various techniques for detecting password leakage. Machine learning, statistical analysis, user profiling, and behavioral biometrics-based approaches can all be used to detect password leakage and protect sensitive information from unauthorized access.

III. PROPOSED SYSTEM

A proposed system for data leakage detection using graphical password with cryptography and steganography can include the following components:

1. Graphical Password Authentication System: A graphical password authentication system can be used to authenticate users based on the images or patterns they select. This system can be designed to store user passwords in an encrypted form to protect them from unauthorized access.
2. Cryptographic Techniques: Cryptographic techniques such as encryption and decryption can be used to secure user passwords and prevent unauthorized access. Passwords can be encrypted using symmetric or asymmetric encryption algorithms and stored in a secure database.
3. Steganography Techniques: Steganography techniques can be used to hide the encrypted passwords within other images or files to prevent detection. The steganography technique can be applied to the images that are used as part of

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12607

the graphical password authentication system, which can make it difficult for an attacker to identify which images are being used for authentication.

4. Data Leakage Detection Mechanism: A data leakage detection mechanism can be used to identify any unauthorized access to the graphical password authentication system. This mechanism can be designed to detect anomalies in user behavior, such as the sequence of images used or the time it takes to enter a password. Machine learning algorithms, statistical analysis, or user profiling techniques can be used for data leakage detection.

5. Notification and Response Mechanism: A notification and response mechanism can be implemented to alert system administrators or users of any potential security breaches. Once a potential security breach is detected, the system can notify the appropriate personnel to investigate and take corrective actions.

By combining cryptographic and steganography techniques with a graphical password authentication system and a data leakage detection mechanism, this proposed system can provide enhanced security against data leakage. The use of steganography can make it more difficult for attackers to identify the images used in the graphical password authentication system, while the use of cryptographic techniques can protect the passwords from unauthorized access. The data leakage detection mechanism can detect any potential security breaches, and the notification and response mechanism can alert the appropriate personnel to take corrective actions.

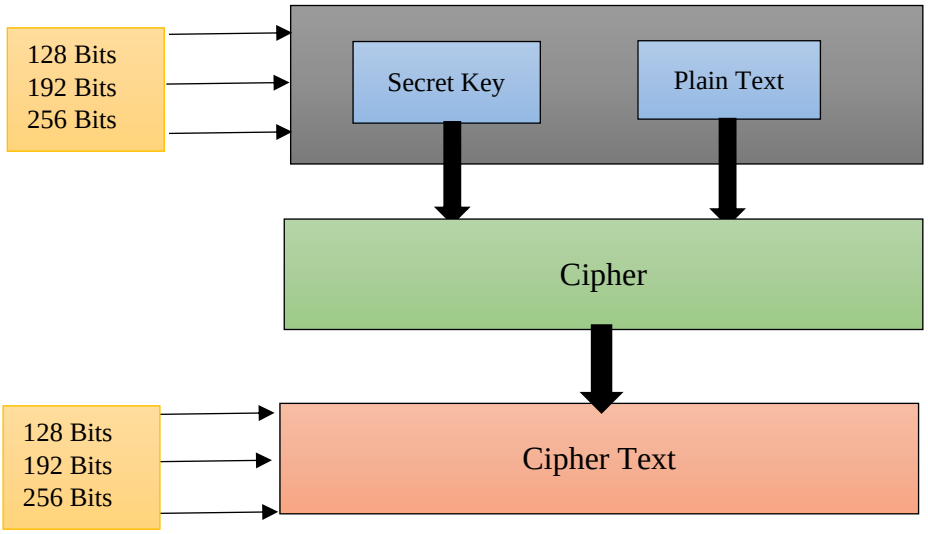

Fig. 1: Project Flow

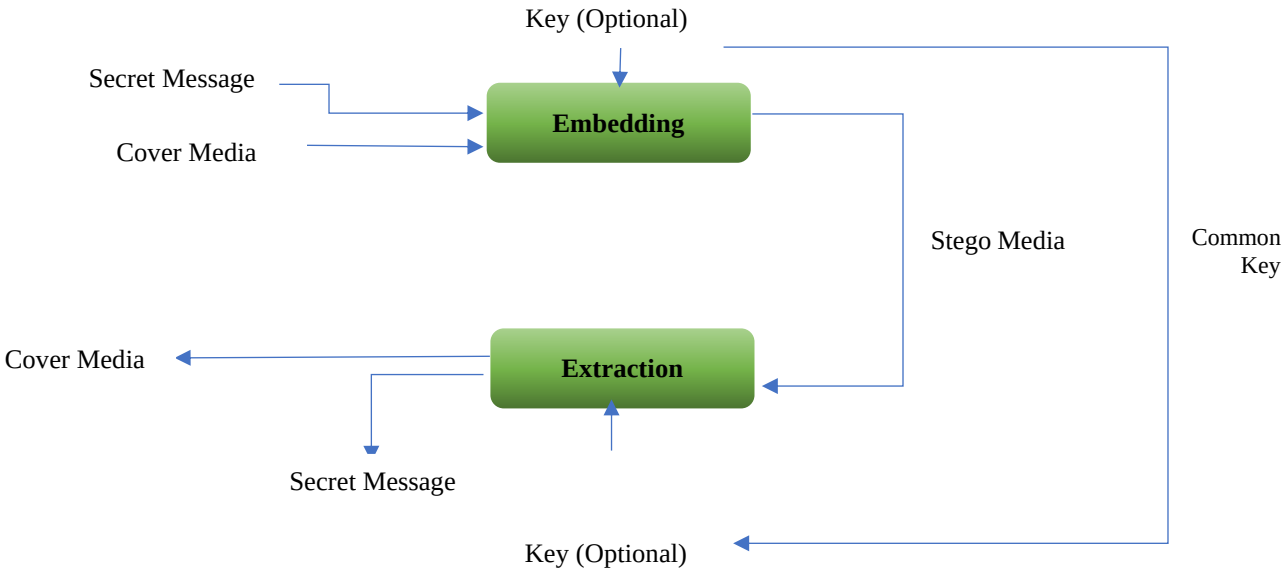

Fig. 2: Project Flow Diagram

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12607

Steganography is the art and science of hiding information within other information to keep it secret from unauthorized parties. The process of steganography typically involves the following steps:

1. Select the Cover Medium: The cover medium is the medium that will be used to hide the secret message. It can be any digital file, such as an image, audio file, video file, or text file.
2. Encode the Secret Message: The secret message is encoded using a steganography algorithm to make it invisible or difficult to detect. The message is typically encoded using digital techniques such as LSB (Least Significant Bit) or DCT (Discrete Cosine Transform).
3. Embed the Secret Message: The encoded secret message is then embedded into the cover medium using the steganography algorithm. The message is typically embedded in a way that does not change the appearance or quality of the cover medium.
4. Transmission or Storage: The cover medium with the embedded secret message can then be transmitted or stored, and it will appear to be an ordinary cover medium without any indication that it contains a hidden message.
5. Extraction of Secret Message: To extract the secret message from the cover medium, the steganography algorithm is used to decode the message from the cover medium.
6. Steganalysis: Steganalysis is the process of detecting the presence of hidden information within a cover medium. It involves analyzing the cover medium to detect any anomalies or changes that may indicate the presence of hidden information. Steganalysis can be performed using various techniques, such as statistical analysis, machine learning, and visual inspection.
7. Key Management: The security of steganography techniques depends on the secrecy and protection of the encryption key. Key management involves the secure storage and sharing of encryption keys to ensure that only authorized parties can access the hidden information.
8. Quality Control: Quality control is an important aspect of steganography. The embedding process should be carefully controlled to ensure that the cover medium remains of high quality and the embedded message is not detectable. Quality control techniques can include visual inspection, objective measures of quality, and human perception studies.
9. Steganography Tools: There are many steganography tools available that automate the steganography process. These tools can help users encode and embed messages into cover media and extract hidden messages from them. However, it is important to use trusted and secure tools to ensure the confidentiality of the messages.
10. Legal and Ethical Considerations: Steganography can be used for both legal and illegal purposes. It is important to consider the legal and ethical implications of using steganography, particularly when it comes to the privacy and confidentiality of personal information. The use of steganography should always comply with applicable laws and ethical standards.

It is important to note that steganography is not a replacement for encryption, but rather a complementary technique. Encryption is used to protect the confidentiality of the message, while steganography is used to hide the existence of the message. By combining encryption and steganography, messages can be protected and hidden from unauthorized parties.

- *Cryptography*

It is the science of using mathematical algorithms to protect information by encrypting it in a way that makes it difficult or impossible for unauthorized parties to access or read it. Cryptography is used to provide confidentiality, integrity, authentication, and non-repudiation of data.

The process of cryptography typically involves the following steps:

1. Encryption: Encryption is the process of transforming the original plaintext message into an unintelligible ciphertext message using a cryptographic algorithm and a secret key. The resulting ciphertext message is meaningless and unreadable without the key.
2. Decryption: Decryption is the process of transforming the ciphertext message back into the original plaintext message using the same cryptographic algorithm and the secret key that was used for encryption.
3. Key Management: Key management involves the secure storage and sharing of cryptographic keys to ensure that only authorized parties can access the protected information. Key management also involves the creation and distribution of new keys on a regular basis to ensure the ongoing security of the system.
4. Authentication: Authentication is the process of verifying the identity of a user or device. Cryptography can be used to provide authentication by using digital signatures or message authentication codes (MACs) to verify the integrity and authenticity of a message.
5. Non-repudiation: non-repudiation is the ability to prove that a message was sent by a specific sender and that the sender cannot deny sending the message. Cryptography can be used to provide non-repudiation by using digital signatures to prove the identity of the sender.

DOI: 10.17148/IJARCCE.2023.12607

There are various cryptographic algorithms that can be used for encryption and decryption, such as AES, RSA, and SHA. The choice of algorithm depends on the specific requirements of the system, including the level of security needed, the size of the plaintext message, and the processing power available.

Overall, cryptography plays a critical role in securing information and protecting privacy. It is used in a wide range of applications, including online banking, e-commerce, messaging, and email, and is essential for ensuring the confidentiality and integrity of sensitive data.

- *AES Algorithm.*

AES (Advanced Encryption Standard) is a symmetric-key encryption algorithm used for encrypting and decrypting electronic data. It was established by the National Institute of Standards and Technology (NIST) in 2001 and is widely used in various applications, including online banking, email, and secure communications.

The AES algorithm uses a symmetric-key approach, which means that the same key is used for both encryption and decryption of data. The key length for AES can be 128 bits, 192 bits, or 256 bits, and the longer the key length, the more secure the encryption.

The AES algorithm operates on a fixed block size of 128 bits and uses a substitution-permutation network (SPN) structure to perform encryption and decryption. The SPN structure is composed of multiple rounds, and each round consists of four operations: SubBytes, ShiftRows, MixColumns, and AddRoundKey.

The SubBytes operation substitutes each byte in the input with a corresponding value from a fixed S-box. The ShiftRows operation shifts the bytes in each row of the input by a certain number of bytes. The MixColumns operation performs a matrix multiplication on each column of the input. The AddRoundKey operation performs an XOR operation between the input and a round key derived from the encryption key.

The AES algorithm is known for its high security and efficiency, and it has been adopted as the standard encryption algorithm by various organizations, including the U.S. government. It has also been extensively studied by the cryptographic community, and no significant vulnerabilities have been found in the algorithm.

Overall, AES is a widely used and trusted encryption algorithm that provides strong security and efficient performance for protecting electronic data.

There are several widely-used cryptographic algorithms that are considered secure and effective, including:

a) AES (Advanced Encryption Standard): AES is a symmetric-key encryption algorithm that uses a fixed block size of 128 bits and can have a key length of 128 bits, 192 bits, or 256 bits. AES is widely used and trusted for its strong security and efficient performance.

b) RSA (Rivest-Shamir-Adleman): RSA is a public-key encryption algorithm that is widely used for secure communications, digital signatures, and other applications. RSA uses a variable-length key for encryption and decryption.

c) ECC (Elliptic Curve Cryptography): ECC is a public-key cryptography algorithm that is used for secure communications and other applications. ECC uses shorter keys than RSA, which makes it more efficient for some applications.

d) Blowfish: Blowfish is a symmetric-key encryption algorithm that uses a variable-length key and a variable number of encryption rounds. Blowfish is known for its efficiency and has been widely used in various applications.

e) SHA (Secure Hash Algorithm): SHA is a family of cryptographic hash functions that are widely used for data integrity and other applications. SHA produces a fixed-length message digest that can be used to verify the integrity of data.

f)

Overall, the choice of algorithm depends on the specific requirements of the application, and it is important to choose an algorithm that provides the appropriate level of security and efficiency for the task at hand.

IV. CONCLUSION AND FUTURE WORK

In conclusion, data leakage detection using graphical password with the combination of cryptography and steganography is a promising approach for improving data security in various applications. The proposed system provides an additional layer of security by using a graphical password, which is less vulnerable to attacks than traditional password-based authentication methods. Moreover, the use of cryptography and steganography enhances the security of the system by providing encryption of sensitive data and hiding it within images or other media. This makes it more difficult for

DOI: 10.17148/IJARCCE.2023.12607

attackers to intercept and decode the data, even if they manage to gain access to it. Overall, the proposed system offers a robust and secure solution for detecting and preventing data leakage in various applications, including online banking, e-commerce, and social networking. It provides an effective mechanism for protecting sensitive data and preserving the confidentiality and integrity of the information. Further research can be done to improve the proposed system and to evaluate its performance under various conditions.

There are several potential areas for future work related to data leakage detection using graphical password. Some of these include:

a) Enhancing the graphical password scheme: Future work can focus on improving the graphical password scheme by exploring different types of graphical passwords, such as using multiple images or using a combination of images and text. This can make the password scheme more robust and less vulnerable to attacks.

b) Developing new steganographic techniques: Steganography is an evolving field, and new techniques are constantly being developed. Future work can explore the use of novel steganographic techniques for hiding data within images or other media to enhance the security of the proposed system.

c) Evaluating performance under different conditions: The proposed system needs to be evaluated under different conditions, such as varying network speeds and different types of attacks, to determine its effectiveness in real-world scenarios.

d) Integrating with other security measures: Future work can explore the integration of the proposed system with other security measures, such as intrusion detection systems, to provide a more comprehensive security solution.

e) Scaling up the system: The proposed system needs to be scaled up to handle large volumes of data in real-world applications. Future work can focus on developing efficient algorithms and techniques to handle large-scale data leakage detection using graphical password.

Overall, future work can focus on enhancing the proposed system to provide a more robust and secure solution for detecting and preventing data leakage in various applications.

REFERENCES

[1]. **"A Machine Learning-Based Graphical Password System for Detecting Insider Threats"** by J. Zhu, Y. Zuo, Y. Ren, and X. Liu (2020). This paper proposes a machine learning-based graphical password system that can detect insider threats by analyzing user behavior patterns.

[2]. **"A Novel Data Leakage Detection System Based on Graphical Passwords" by Y. Li and X. Wu (2020).** This paper presents a novel data leakage detection system that uses graphical passwords as an authentication method. The system can detect data leakage in real-time by monitoring user behavior.

[3]. **"A Comparative Analysis of Graphical Password Schemes for Data Leakage Detection" by S. Kumar, S. Singh, and R. N. Yadav (2019).** This paper compares different graphical password schemes for data leakage detection and evaluates their performance in terms of accuracy, efficiency, and security.

[4]. **"A Novel Approach for Data Leakage Detection Using Graphical Password" by N. A. Khan and A. A. A. Bakar (2018).** This paper proposes a novel approach for data leakage detection that uses graphical passwords as an authentication method. The approach is based on the analysis of user behavior patterns and can detect data leakage in real-time.

[5]. **"Data Leakage Detection in Cloud Computing Using Graphical Password Authentication" by R. P. Srivastava, S. K. Gupta, and S. K. Singh (2018).** This paper proposes a data leakage detection system for cloud computing that uses graphical passwords as an authentication method. The system can detect data leakage by monitoring user behavior patterns in real-time.

[6]. **"Behavioral Biometrics for Data Leakage Detection in Graphical Password Systems" by M. A. Muñoz, J. A. Ortega-García, and J. Fierrez (2021).** This paper proposes a data leakage detection approach using behavioral biometrics in graphical password systems. The approach uses machine learning techniques to analyze user behavior patterns and detect abnormal behavior that may indicate data leakage.

[7]. **"A Survey of Graphical Password Schemes for Data Leakage Detection" by S. S. Rao, R. S. Rao, and S. K. Pradhan (2020).** This paper provides a survey of graphical password schemes for data leakage detection, including their advantages, disadvantages, and limitations.

[8]. **"An Efficient Data Leakage Detection Technique using Graphical Passwords and User Behavior Analysis" by M. R. Chowdhury, M. H. Kabir, and M. S. Islam (2020).** This paper proposes an efficient data leakage detection technique that uses graphical passwords and user behavior analysis. The technique is based on the analysis of user behavior patterns and can detect data leakage in real-time.

[9]. **"A Novel Data Leakage Detection Approach using Graphical Passwords and Keystroke Dynamics" by M. Alam and M. T. Hossain (2019).** This paper proposes a novel data leakage detection approach using graphical

DOI: 10.17148/IJARCCE.2023.12607

passwords and keystroke dynamics. The approach uses machine learning techniques to analyze user behavior patterns and detect abnormal behavior that may indicate data leakage.

[10]. **"Data Leakage Detection using Graphical Passwords: A Survey" by P. S. Varghese and J. K. Abraham (2019).** This paper provides a survey of data leakage detection approaches using graphical passwords, including their advantages, disadvantages, and limitations.

[11]. **"A Framework for Continuous Data Leakage Detection Using Graphical Passwords and Eye-Tracking" by E. Oyibo, D. Wolff, and M. Junker (2018).** This paper proposes a framework for continuous data leakage detection that uses graphical passwords and eye-tracking. The framework can detect data leakage in real-time by monitoring user behavior patterns and eye movements.

[12]. **"Graphical Password Based Data Leakage Detection using Machine Learning" by P. Sharma and P. Gupta (2018).** This paper proposes a graphical password-based data leakage detection approach that uses machine learning techniques to analyze user behavior patterns. The approach can detect data leakage in real-time and provide early warning notifications to prevent data loss.

