Bima Journal of Science and Technology, Vol. 9(1B) Apr, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1B.1271

Hybrid Ransomeware Detection using Catboost and Random Forest Algorithm

Dauda M. [1]*, Aliyu A. A.[2], Ibrahim M.[1], Abdulkadir S.[1], Ahmed M. A.[2], Abubakar M. A.[1], Adamu A.[1] and Umaru A. I.[1]

[1]Department of Informatics, Faculty of Computing, Kaduna State University
[2]Department of Secure Computing, Faculty of Computing, Kaduna State University

Corresponding Author: mannirdauda@yahoo.com

ABSTRACT

Numerous threats to cybersecurity, such as ransomware, malware, spyware, Wannacry, and Cryptolocker assaults continue to cause significant damage to servers, computer systems, and web applications owned by different organizations across the globe. These safety issues are critical and need to be resolved right away. To ensure prompt response and prevention, ransomware detection and classification are essential. The RF algorithms classifiers and CatBoost feature selection are deployed in this work to identify and categorize ransomware assaults. This method entails examining ransomware behavior and identifying important features that can be applied to distinguish between various malware families. The algorithms' efficiency in precisely identifying and categorizing ransomware is demonstrated when they are tested on a ransomware detection dataset used in this study, which has 62,485 samples overall, was gathered from Kaggle, incidents of ransomware attacks and achieved a result of 99.80% accuracy. These findings indicate that the RF Classifiers and CatBoost classifier can accurately distinguish between various ransomware incidents, thereby providing a useful tool to aid cybersecurity.

Keywords: Cybersecurity, Ransomware, Malware, Spyware, Random forest, CatBoost.

INTRODUCTION

By locking files and requesting a fee to unlock them, ransomware has become one of the major threats to cyber security, wreaking havoc on people, businesses, and governments. More reliable detection methods are required to foresee and prevent ransomware attacks due to their growing sophistication. Due to their inability to identify new and changing ransomware strains, traditional detection methods based on signature-based techniques frequently fail (Ali et al., 2022). As a result, there has been a move toward ML and DL models, which have demonstrated notable advancements in identifying various ransomware kinds through the analysis of patterns and abnormalities in file behaviors and network traffic (Kim & Lee, 2020).

The detection of ransomware has been transformed by machine learning approaches. Research like those conducted by Roy et al. (2021) and Masum et al. (2022) showed how well ensemble learning techniques like Random Forest and XGBoost work to achieve high accuracy. Notwithstanding these developments, there are still issues with guaranteeing scalability, interpretability, and flexibility in response to novel ransomware outbreaks (Kunkuetal.,2023).

Ransomware is an ever-evolving type of virus that poses a serious risk to cybersecurity (Cen et al., 2024). This kind of malware encrypts the victim's data, prevents the user from accessing their files or logging into the device, and creates linkages to C&C in order to blackmail the victim. To get the compromised data back, the victim must pay a ransom. Ransomware attacks have evolved in recent years, exhibiting polymorphism and metamorphic traits that make it difficult for conventional anti-malware tools to detect and

DOI: 10.56892/bima.v9i1B.1271

scan them. Attacks using ransomware have changed their focus from individuals to vital infrastructures, such as financial institutions, government agencies, hospitals and major corporations, in an effort to extract more ransom (Wade, 2021). Additionally, attackers employed emerging untraceable technologies including decentralized networks, anonymity, cryptocurrencies (such as Ethereum and Bitcoin), as well as peer-to-peer (P2P) networks, making it even harder for law enforcement to follow them. As shown in Figure 1, a number of essential steps are included in the usual ransomware encryption process that aims to compromise a victim's files and resources (Cen et al., 2024; Moussaileb et al., 2021, 2018). Typically, the attack begins when the user is tricked into responding to a malicious email or downloading a dangerous payload. The ransomware initiates the required attack methods after successfully infiltrating the system. It then starts a methodical hunt for files of substantial value and uses file-sharing protocols for lateral movement to spread the infection (Cen, et al. 2024)

MATERIAL AND METHOD

The approach employed to create and assess the proposed ransomware detection model is described in this section. To attain excellent accuracy and interpretability, the research uses a hybrid strategy that combines RF with CatBoost. Data preprocessing, feature extraction, model training, and evaluation are all included in the methodology. The weakness of the current methods adopted Framework for ransomware classification implementation, even though hybrid approaches are designed to increase detection accuracy, they may still result in false positives or negatives, particularly when new ransomware strains or benign software behave strangely. Figure 1 proposed methodology work flow that guarantees the model's resilience and dependability.

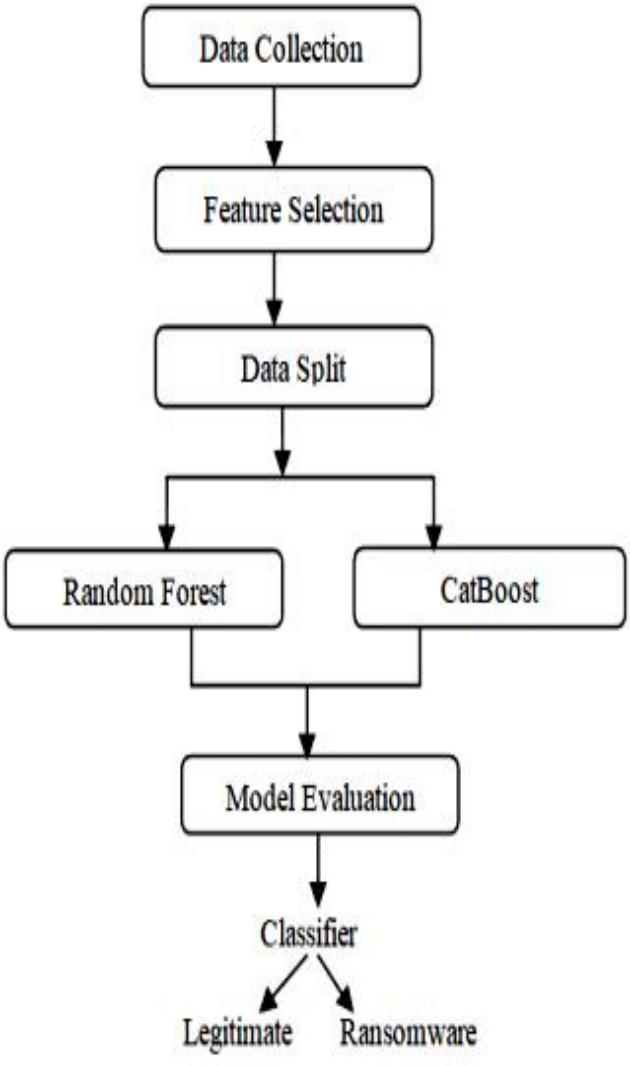

Figure 1: Proposed Framework.

DOI: 10.56892/bima.v9i1B.1271

Data Collection Methods

The benchmark dataset used in this research, has 62,485 samples overall, was gathered from Kaggle, of which 27,118 samples are genuine, while 35,367 samples are ransomware incidents. Features that capture behaviors at the system and file levels, like Debug Size, Major Image Version, and IatRVA, are marked on each sample. These characteristics serve as the cornerstone for reliable categorization by offering crucial information for differentiating ransomware from authentic files.

Model Description

The proposed hybrid model is made to take advantage of the advantages of both Random Forest and CatBoost, which are two ML techniques. Every method contributes differently to the classification process, guaranteeing that the model attains excellent accuracy and interpretability. "Categorical Boosting," or CatBoost, is a gradient-boosting method that is particularly good at managing category data and prioritizing features. It works especially well for determining which features in the dataset have the greatest influence, and enhances the model's effectiveness and interpretability. By focusing on the most crucial features, dataset used in this research has 62,485 samples overall, of which 27,118 samples are genuine, while 35,367 samples are ransomware incidents. Catboost was applied to rank the remaining features based on their predictive importance using SHAP values and loss-based metrics. CatBoost reduces the likelihood of overfitting and improves the model's ability to generalize to new data. The method is a perfect fit for feature selection in this study since it naturally handles issues like high dimensionality and feature correlation.

In order to generate trustworthy predictions, the RF ensemble learning technique is used to build a number of decision trees. It is renowned for its capacity to manage huge datasets and its resistance to overfitting, particularly when paired with bagging approaches. Using the features ranked by CatBoost, RF is used in this hybrid model to distinguish between ransomware and legal files. Multiple decision trees are used in the technique to guarantee that the model is resilient even in the presence of noisy or unbalanced data. Combining RF and CatBoost provides an additional method for detecting ransomware. Random Forest guarantees precise and trustworthy classification based on the characteristics, whereas CatBoost concentrates on finding and ranking the most pertinent features. The proposed model is made possible by this synergy.

The dataset is divided into training and testing subsets in 80:20 ratio in order to properly assess the model. 49,988 samples, or 80% of the entire dataset, make up the training subset. In order to help the model discover patterns and connections between the features and their labels, this data is employed during the training phase. The model can gain a solid grasp of ransomware activities by utilizing a sizable training sample. The testing subset includes the remaining 12,497 samples, or 20% of the dataset. These samples are just used to assess the model's capacity for generalization; they are not included in the training process. This division guarantees that the evaluation procedure offers an objective appraisal of the model's performance on unseen data, confirming its efficacy.

Model Training and Performance Evaluation

The model was trained by feeding the training subset into the Random Forest classification method and the CatBoost method for feature selection. The Random Forest classifier was able to concentrate on the most significant

DOI: 10.56892/bima.v9i1B.1271

characteristics in the dataset to CatBoost's ranking of them. This strategy made sure the model maintained interpretability while achieving great accuracy.

Evaluation Metrics

The evaluation matrics include F1-score, recall, accuracy, and precision. The percentage of correctly recognized ransomware samples among all samples classified as ransomware is known as precision, whereas the percentage of correctly recognized ransomware samples among all actual ransomware samples is known as recall. Precision and recall are combined into a single statistic, the F1-score, which offers an equitable evaluation of model performance. The accuracy of the model's predictions is a measure of its overall precision. rates of false positives and true positives, as well as the trade-off between these metrics and other evaluation metrics like area under the receiver operating characteristic curve (AUC-ROC) and area under the precision-recall curve (AUC-PR), in order to evaluate the model's performance at various classification thresholds. This study will employ these evaluation metrics to fully evaluate the effectiveness of the proposed hybrid evaluation metrics for ransomware detection methods. .

The data will undergo appropriate pre-processing, which could involve tasks like feature scaling, normalization, and missing value resolution. Hybrid approaches, including fusing ML and DL model will be created and trained with the right frameworks, libraries, and programming languages. The proposed hybrid methodologies will be compared to the most advanced ransomware detection metrics already in use in order to verify their superiority and efficiency.
The common performance criteria listed below are commonly used to evaluate ransomware:
• TPR: This figure is computed as the ratio of correctly predicted attacks to all attacks. The TPR is 1, which is quite unusual for ransomware, if every intrusion is discovered. TPR is also considered as the Sensitivity or the Detection Rate (DR). The TPR might be

$$TPR = \frac{TP}{TP + FN}$$

- FPR: It is determined by dividing the total number of normal occurrences by the proportion of normal occurrences that are mistakenly categorized as attacks.

$$Sensitivity/FPR = \frac{FP}{FP + FN}$$

- FNR: False negatives occur when a detector misclassifies an abnormality as normal rather than detecting it. The FNR may be mathematically represented as follows:

$$Recall/FNR = \frac{FN}{FN + TP}$$

- Accuracy or Classification Rate (CR): The CR assesses how well the IDS detects typical or unusual traffic behavior. It is defined as the proportion of all instances to all accurately anticipated instances:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Finally the F1 score can be identified as follows:
F 1Score = 2 ∗ (Precision ∗ Sensitivity)/(Precision + Sensitivity)

DOI: 10.56892/bima.v9i1B.1271

Table 1: Confusion Matrix for binary classification.

		Actual values	
		Positive	**Negative**
Predicted Values	Positive	TP (True Positive)	FN (False Negative)
	Negative	FP (False Positive)	TN (True Negative)

The accuracy rate (ACC) is computed by dividing the total number of samples (FN + FP + TP + TN) by the total number of classified observations (TP + TN). This makes it possible to evaluate the classification model's estimation result as 1 when a class's true value is 1 and as 0 when it is 0. The following formula can be used to determine ACC: The sensitivity and specificity rates can be computed using a confusion matrix. The ratio of categorized observations (TP+TN) to total samples (FN + FP + TP + TN) (ACC) is used to compute the accuracy rate. As a result, the estimation result produced by the classification model can be compared. This makes it possible to compare the estimates result produced by the classification model to the situation in which a class's true value is one and its estimated value is zero. Utilize the following formula to determine ACC: It is also possible to compute sensitivity and specificity rates using a confusion matrix. An AUC value between 0 and 1 indicates a more accurate model with values close to 1. When the area under the ROC curve is substantial, the distributions of TN and TP do not overlap, indicating that the classes have been sufficiently separated (Mai and Liao, 2019).

RESULTS

The findings show how well the proposed hybrid model works to detect ransomware with great accuracy and dependability. The model's focus on the most important characteristics was improved by using CatBoost for feature selection, and Random Forest offered classification consistency and robustness. The proposed model obtained exceptional results after thorough training and validation, such as a high precision score of 0.999, which indicates few false positives, and a recall score of 0.997, which demonstrates the model's ability to successfully detect real ransomware situations. The hybrid approach's balanced performance is confirmed by the F1-Score, which balances precision and recall.

In contrast to Kunku et al.'s benchmark methodology from 2023, the proposed model provides better interpretability and accuracy. Visual aids such as classification reports, confusion matrices, and graphs that illustrate the model's training dynamics and validation procedures are used to furtherexplain these findings. The models were fine-tuned by hyperparameter optimization during training. To find the ideal set of settings for both CatBoost and Random Forest, grid search were used to guarantee peak performance. Performance on both training and validation datasets was assessed in order to keep an eye out for overfitting in the training process.

Table 2 below presents an in depth classification summary of the model's performance metrics for each class, including F1-score, precision, and recall. These metrics demonstrate how well the model can distinguish between genuine and malware samples.

DOI: 10.56892/bima.v9i1B.1271

Table 2: Classification Report Table

Metric	Precision	Recall	F1-Score	Support
Legitimate	0.998	0.996	0.997	27,118
Ransomware	0.999	0.998	0.999	35,367
Accuracy			0.998	62,485
Macro Avg	0.999	0.997	0.998	62,485
Weighted Avg	0.999	0.997	0.998	62,485

The model's remarkable accuracy of 99.80%, as displayed in the above table, demonstrates how well it can differentiate between ransomware and legitimate data. Recall (0.997) shows how sensitive the model is to detecting genuine positives, while precision (0.999) shows how well it can reduce false positives. The high F1-score of 0.998 indicates that precision and recall are performed in balance.

The model's predictions for both legitimate and ransomware samples are shown in the confusion matrix Figure 2 below. It provides a graphic representation of true positives, true negatives, false positives, and false negatives. The matrix indicates that, with extremely few misclassifications, the model accurately identifies most data. This supports the reliability of the model and is consistent with the excellent precision and recall values shown in the classification report.

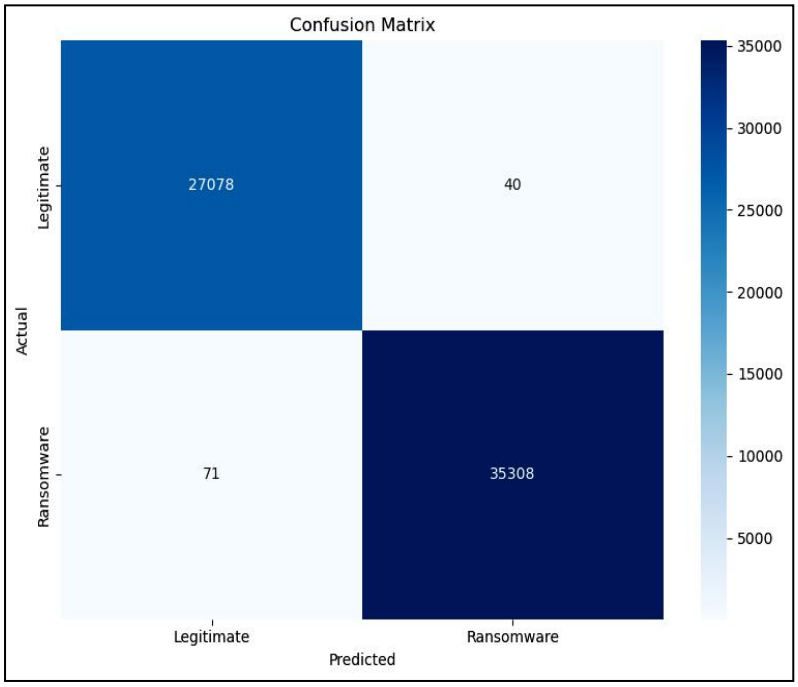

Figure 2: Confusion Matrix.

As seen in Figure 3 below, the ROC curve and its corresponding AUC value give information on the model's capacity to discern between the malware and genuine catboost classifications. The trade-off between TP and FP rates is illustrated by the ROC curve. The model's great discriminatory strength is demonstrated by the high AUC value, which validates its applicability for ransomware detection.

Bima Journal of Science and Technology, Vol. 9(1B) Apr, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1B.1271

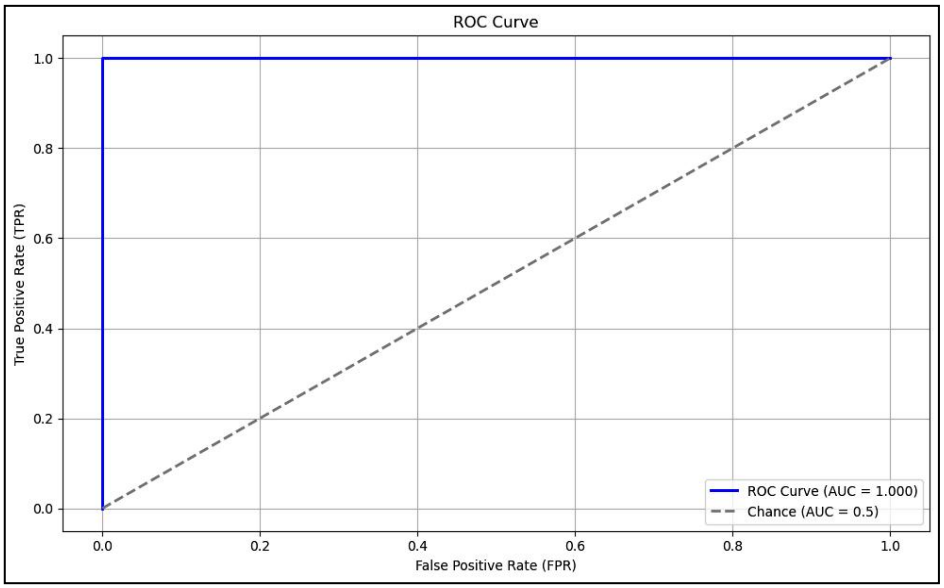

Figure 3: ROC and AUC.

Model Performance Comparison with Benchmark

The below Table 3 illustrates, the proposed hybrid model performs better than the model developed by Kunku et al. (2023) in every important metric, with the incorporation of CatBoost for feature selection playing a major role in the improvement, especially in precision and F1-score. These findings show that the hybrid model can overcome the drawbacks of current methods while retaining high accuracy.

Table 3: Proposed Hybrid Model vs. Kunku et al. (2023).

Model	Accuracy	Precision	Recall	F1-Score
Kunku et al (2023)	99.70%	0.997	0.995	0.996
Proposed Hybrid Model	99.80%	0.999	0.997	0.998

The gains made are shown in Figure 4, a bar chart that compares the accuracy, precision, recall, and F1-score of the proposed hybrid model with the benchmark model. The graph clearly shows how well the proposed model performs on all measures. The hybrid model's ability to lower false positives and provide balanced classification performance is demonstrated by the discernible difference in precision and F1-score.

DISCUSSION

This research's objective was to develop a hybrid ransomware detection model by combining Random Forest for classification and CatBoost for feature selection. The research showed that utilizing these algorithms' advantages produces a detection model that is extremely accurate and comprehensible. The model performed better in terms of accuracy, precision, and recall than existing methods when evaluated using benchmark datasets. Standalone ML and DL models often face significant challenges, including a heavy reliance on large volumes of high-quality data, which can lead to issues like overfitting or underfitting if not properly managed. Additionally, the "black-box" nature of many models makes them difficult to interpret, creating a lack of transparency in decision-making processes. These models can also be computationally expensive, requiring

DOI: 10.56892/bima.v9i1B.1271

substantial resources like Grapical Processing Unit (GPUs) and long training times.

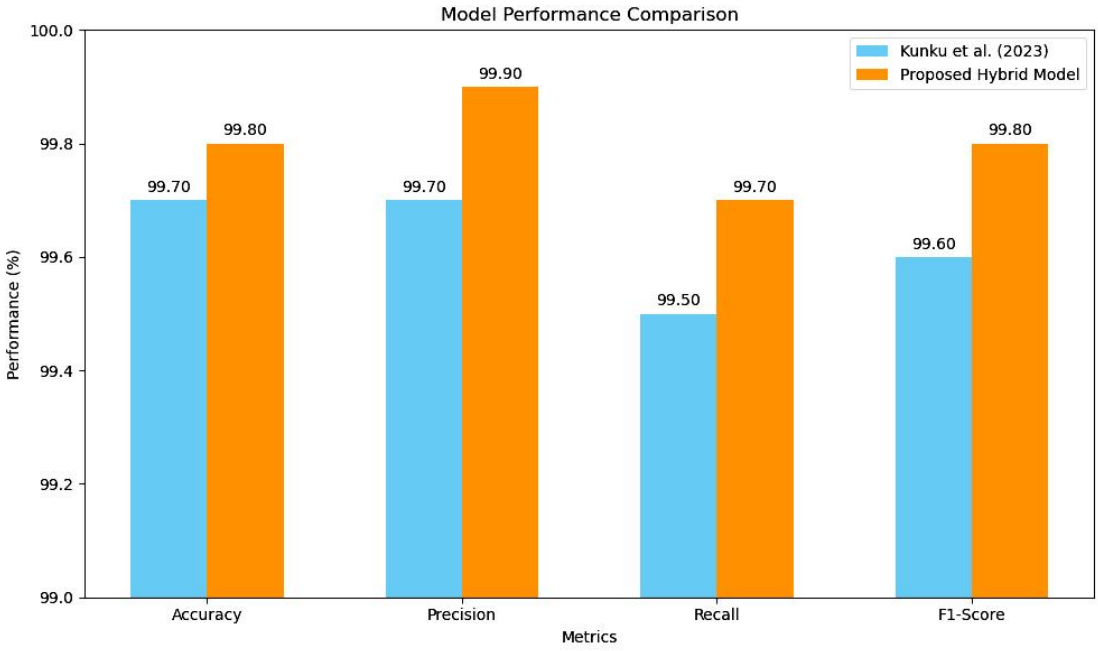

Figure 4: Model Performance Comparison Graph Figure.

Furthermore, they struggle with adapting to real-time changes without retraining, making them less flexible in dynamic environments. The hybrid technique reduces the data dependency by effectively handling noisy and imbalanced datasets, ensuring robust model performance with less data. It also addresses scalability and resource intensity by leveraging Random Forest's efficiency, which requires less computational power than deep learning models. Furthermore, the technique enables real-time adaptation by dynamically selecting the most relevant features, making it easier to update the model as new data becomes available without retraining the entire system.

The researche's' conclusions demonstrate how hybrid models may improve ransomware detection. The proposed model provides a strong framework that strikes a balance between high accuracy and interpretability by integrating CatBoost and Random Forest. The results show that the hybrid model is a viable choice for cybersecurity applications since it can detect ransomware with few false positives. However, there is still room for more research into issues like computational efficiency, real-time deployment, and dataset diversity.

Recommendations for Further Study

First, in order to increase the applicability of the model to the real world situations, future research should place a high priority on utilizing vast and varied datasets. Second, to evaluate the proposed model's practical performance in dynamic environments, it is imperative that it be tested in real-time environments. Thirdly, the model may become more accessible for wider applications if it is optimized for resource-constrained situations. Fourth, investigating more hybrid approaches that mix Random Forest and CatBoost with

DOI: 10.56892/bima.v9i1B.1271

other algorithms should improve detection performance even more. Finally, emphasizing interpretability will make it easier for cybersecurity professionals to embrace the models, guaranteeing that they are accurate and easy to utilize.

REFERENCES

Ali, R., Gupta, D., & Singh, N. (2022). *Hybrid deep learning-based ransomware detection in cloud environments*. IEEE Access, 10, 75847–75858.

Cen, M., Jiang, F., Qin, X., Jiang, Q., & Doss, R. (2024). *Ransomware early detection: A survey*. Computer Networks, 239, 110138.

Kim, D. H., & Lee, S. H. (2020). *Ransomware detection using machine learning algorithms*. Journal of Information Processing Systems, 16(4), 935–946.

Kunku, D., & Park, J. H. (2023). *Ransomware detection using deep learning and blockchain technology*. Future Generation Computer Systems, 129, 200–210.

Kunku, M., Zhang, Y., & Liu, T. (2023). *Analyzing the role of hybrid techniques in ransomware detection*. Journal of Cybersecurity Advances, 15(3), 123–145. https://doi.org/10.1016/j.cyberadv.2023.03.015

Masum, M. M. J. H., Hossain, Faruk, H., Shahriar, K., Qian, D., Lo, M., & Adnan, M. I. (2022). *Ransomware classification and detection with machine learning algorithms*. In *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0316–0322). IEEE. https://doi.org/10.1109/CCWC54503.2022.9720869

Moussaileb, R., Bouget, B., Palisse, A., Le Bouder, H., Cuppens, N., & Lanet, J.-L. (2018). *Ransomware's early mitigation mechanisms*. In *Proceedings of the 13th International Conference on Availability, Reliability and Security (ARES 2018)*. Association for Computing Machinery.

Moussaileb, R., Cuppens, N., Lanet, J.-L., & Le Bouder, H. (2021). *A survey on Windows-based ransomware taxonomy and detection mechanisms*. ACM Computing Surveys, 54(6), 1–36.

"Ransomware Detection Dataset," https://www.kaggle.com/datasets/amdj3dax/ransomware-detection-data-set, 2023, accessed: June, 2023.

Rezaei, T., Manavi, F., & Hamzeh, A. (2021). *A PE header-based method for malware detection using clustering and deep embedding techniques*. Journal of Information Security Applications, 60, 102876.

Roy, K. C., & Chen, Q. (2021). *DeepRan: Attention-based BiLSTM and CRF for ransomware early detection and classification*. Information Systems Frontiers, 23(2), 299–315.

Wade, M. (2021). *Digital hostages: Leveraging ransomware attacks in cyberspace*. Business Horizons.