INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Exploring HPC and ML Techniques on Various Classical, Machine and Modern Ciphers

Venu Nalla[1,2], M Anil Kumar[2], V Samhith[2], G Padmavathi[2]*

1 Department of Computer Science & Engineering, Acharya Nagarjuna University, Guntur
2 Department of Computer Science & Engineering, CRRao AIMSCS, Hyderabad, India

OPEN ACCESS

Received: 08-08-2024
Accepted: 08-05-2025
Published: 30-05-2025

Editor: Special Issue Editors: Dr. N. Pothanna, Dr.T. Jayashree, Dr. V. Ganesh Kumar

Citation: Nalla V, Kumar MA, Samhith V, Padmavathi G (2025) Exploring HPC and ML Techniques on Various Classical, Machine and Modern Ciphers. Indian Journal of Science and Technology 18(SP1): 19-27. https://doi.org/10.17485/IJST/v18si1.icamada16

* **Corresponding author**.

padmavathi@crraoaimscs.res.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To investigate High Performance Computing (HPC) techniques for communication ciphers. To explore machine cipher classification using encrypted image data and ML for 4G & 5G cipher classification based on ciphertext analysis. **Methods:** Employed MPI and CUDA + MPI models to assess the strength of ciphers, classified machine ciphers with image data, and applied ML for identification communication cipher classification. **Findings:** Communication ciphers are resistant to exhaustive key search attacks using HPC. Applied various classification and clustering models on machine ciphers using encrypted image data collected from online, and got the test accuracy up to 93.9%. ML methodologies for classification of 4G and 5G mobile communication ciphers achieved a test accuracy up to 79.2%. **Novelty:** Strength analysis of various Cryptographic algorithms belonging to Classical, Machine & Mobile communication. Classification of Some machine ciphers & Mobile communication ciphers using HPC solely based on ciphertext analysis contributing novel insights into communication cipher security.

Keywords: High Performance Computing; Clustering; Classification; Mobile Communication Ciphers

1 Introduction

Robust cryptographic solutions are gaining significance due to the growing importance of digital connectivity and the sophistication of cyber threats. Due to HPC's incredible processing power and parallel computing capabilities, researchers and professionals are now able to bypass old computational limits and take on complicated issues. As well stated by Sukhpal Singh Gill et al., HPC systems may complete jobs by allocating concurrent resources and using a variety of resources online, using a pooled set of resources [(1,2)].

Ju-Won Park et al., tried to enhance the gains of HPC systems by working on input/output (I/O) performance degradation as a result of higher storage latency and intricate parallel I/O architecture for data access in distributed storage systems [(3)]. Jiwoo Bang et al., have done cluster evaluation in which the test dataset was split up into 3 clusters: cluster 1 is primarily made up of small jobs with operations on standard I/O units; Cluster 2 is made up of middle-sized parallel jobs that are mostly read-only; and

Cluster 3 is made up of large parallel jobs that are heavily write-intensive. This clustering of the jobs demonstrated that achieving high I/O throughput requires the use of a parallel I/O library like MPI IO as well as substantial number of parallel cores [4]. Ajeet Singh et al., have experimented on cryptanalysis of Mono-alphabetic, Poly-alphabetic & Round-reduced modern SIMON ciphers and also displayed the revolutionary role that neural networks play in helping cryptanalysts identify cipher vulnerabilities by feeding these networks data that highlights inherent flaws together with matching encryption keys [5]. This paper explores the utilization of the MPI model and the combined CUDA + MPI model for accelerating cryptanalysis for crypto-algorithm strength validation, especially for 4G & 5G mobile communication ciphers to test the security of sensitive data transmission.

In addition to parallel computing paradigms, this paper also investigates the realm of ML in the domain of cipher analysis and classification. Ju-Won Park et al., investigated four classification methods and used application-specific I/O patterns to identify different apps with an accuracy of above 90% [3]. Lily Schleider et al., conducted an analysis of the clustering performance by various feature representations, using multiple distance measures and their combinations on the original and feature spaces [6]. S. Ramraj & G. Usha have assessed the effectiveness of the Support Vector Machine (SVM) in categorizing network packets according to the kind of application & the kind of data that is communicated within an application where the F1 score of the SVM classifier for encrypted network packets from Virtual Private Networks (VPN) and the WhatsApp mobile application is 0.9 [7]. This work tried the classification and clustering of machine ciphers using data obtained from online simulators, providing valuable insights into the cryptographic landscape. A significant aspect of this study also involves the application of ML methodologies for the classification of 4G and 5G mobile communication ciphers for 20k plaintext for each selected cipher which has not been done before this work. By analyzing ciphertext data obtained from encrypted communications, predictive models are developed that can discern the underlying cryptographic algorithms employed in modern communication protocols [8].

1.1 Cryptographic Algorithms

Cryptographic algorithms are the backbone of secure communication in the digital age. Table 1 presents the Cryptographic algorithms used in this work.

Table 1. Cryptographic algorithms used in this work

Cipher Type	Algorithm	
Classical Ciphers	Hill	
	SIGABA [10]	
Machine Ciphers	Enigma	
	KL-7	
	Fialka	
	RC4	
Modern Ciphers	Mobile Communication Ciphers	AES
		ZUC [9]
		SNOW

1.2 Cryptanalysis

Cryptanalysis is the study of analyzing and decoding cryptographic systems to find weaknesses and crack the codes to obtain illegal access to protected data. Table 2 shows techniques covered in this work.

2 Methodology

2.1 Cryptanalysis using High Performance Computing

2.1.1 MPI model

The MPI model is the process of using MPI libraries for parallel programming applications, allowing multiple processes to run concurrently and communicate with each other by passing messages. Every process has a memory area, and message forwarding facilitates communication between processes. One of the Strength parameters of cryptographic algorithms (or ciphers) lies in their secret key and current computational power, which they can withstand. The Brute-force attack is a better way of checking

Table 2. Cryptanalysis techniques used in the work

Method			Description
HPC	CUDA		A parallel programming language i.e., PPL that builds upon languages like C, giving developers a comfortable platform for transferring tasks to GPU
	MPI		A programming paradigm to develop parallelly executable programs that give users access to the processing power of numerous CPUs
ML	Clustering techniques [13,14]	K-means	An algorithm that divides data into K clusters iteratively according to similarity by minimizing the centroids of each cluster.
		Hierarchical	Forms a hierarchy of clusters using top-down divisive or bottom-up agglomerative methods.
	Classification techniques	KNN	An approach that designates a class label to an input by a majority vote of its k nearest neighbors.
		Naïve Bayes	Calculates the likelihood of every class based on the data and picks class with the highest probability.
		Logistic Regression	Regression-based model that finds the probability that a specific input is a member of a particular class.
		Decision tree	A tree-based model to divide the data into subgroups recursively.
		Support vector Machine	Supervised learning algorithm that finds which hyperplane in feature space best divides classes.
		Random forest	A technique that builds many decision trees during training and outputs the mean forecast of each tree.
	Features [15]	Gram frequencies	Frequencies of n-grams in a given text or dataset.
		Entropy	It quantifies the average amount of information produced by a random variable.
		Long repeat	It identifies repeated sequences within the ciphertext, which may indicate patterns in the underlying plaintext.
		Index of coincidence	It measures the likelihood that two randomly selected characters in the ciphertext match.
		No. of unique characters	It counts the distinct characters present in the ciphertext.
		Di-graphic IoC	It extends the IoC to pairs of characters, providing a more detailed analysis of ciphertext.
		TF-IDF	TF = No. of occurrences of a word in selected text / total no. of words in selected text IDF = log (total no. of documents / documents containing) TF-IDF = TF * IDF
	Miscellaneous	Elbow method	To figure out a dataset's ideal number of clusters (K) for K-means clustering. [11,12]
		Silhouette score	It measures how similar an object is to its own cluster (cohesion) compared to other clusters (separation) [11]

it. This attack can be performed by effectively dividing the task into embarrassingly independent parts, which is the best case for MPI. When the program initializes MPI, it retrieves the rank and size of the MPI communicator. Then, MPI finalizes after each sub-task returns to the master task to maintain consistency. Non-blocking communication is used, which significantly improves performance in this case.

2.1.2 CUDA + MPI model

A popular method for parallel computing, particularly in high-performance computing (HPC) settings, is to combine CUDA with MPI. The NVIDIA GPU parallel computing platform, programming paradigm CUDA, and distributed computing standard MPI enable communication and synchronization between processes. The work is distributed between MPI processes, distributed data across processes, and CUDA within each method is used to perform parallel computations. When needed,

synchronize data between MPI processes using MPI communication functions. Combining CUDA and MPI requires careful synchronization and data management to ensure the correctness and efficiency of the parallel program.

2.2 Machine Learning-based Cryptanalysis

2.2.1 Data collection / Corpus generation

For machine ciphers data, the Ciphertext in the image format (byte code) is collected from online (https://cryptii.com/). Then, converted image to machine encoded format through optical character recognition by using Pytessecract software. The machine encoded format is then saved as a .csv file.

For mobile communication ciphers data, 20000 ciphertext are collected from each AES, ZUC and Snow ciphers thereby, 60000 in total.

2.2.2 Text Featurization

An emphasis on feature extraction from the ciphertext can aid in analysis because of this randomness. The general features used in classifying ciphers are Entropy-based analysis, Long repeat, Index of coincidence, number of unique characters, Di-graphic IoC, etc. And in machine learning, we often deal with the data sets having one or more levels. Sometimes we need to convert the labels to numerical format so that machine learning algorithms give some better results. One such method is label encoding where, it converts categorial variables to numerical numbers. For example, (low, mid, high) can be converted to (0,1,2).

2.2.3 Clustering

The clustering techniques like K-means and hierarchical clustering methods are applied. Then we see some performance characteristics of the clustering to tell how well the clustering has done.

2.2.4 Classification

Here along with the vector format of text we implement other features for the classification of text. The models like KNN, Logistic Regression, Random Forest, SVM, decision tree and Naive Bayes are used.

3 Results and Discussion

3.1 Cryptanalysis using High Performance Computing

A High-Performance computer with CPU + GPU nodes of 100 TFLOPS of sustained HP LINPACK performance along with 100TB usable high-performance storage is used to evaluate the models' performance in relation to the ciphers.

3.1.1 MPI model on Hill cipher (Classical cipher)

The analysis of Hill cipher is performed using MPI-Programming by generating all probable plain text for matrices of large sizes on a High-Performance Computing Device. A distributed parallel processing application is developed on novel multi-core architecture that uses cipher-text only attack to generate all possible outputs. The work load is distributed using master node to worker nodes.

Table 3. Analysis of HILL cipher using MPI

Matrix size	Cores	Time
2x2	1(single-PC)	4 hours 21 mins
3x3	1(single-PC)	Crashed after 18 hours
4x4	1(single-PC)	Crashed
2x2	4 (Multi-core)	2 hours 17 mins
3x3	4 (Multi-core)	9 hours 3 mins
4x4	4 (Multi-core)	26 hours 13 mins
2x2	15 (Multi-core)	20 mins
3x3	15 (Multi-core)	2 hours 12 mins
4x4	15 (Multi-core)	9 hours

The duty of the master node for this case study is to distribute the overall workload equally to all sub-nodes, send the matrix with which the Key matrix needs to be multiplied (which is all possible matrices by using the given offset value), receive the data after the multiplication and prints it out onto a text file. The Duty of the worker node for this case study is to calculate the determinant of the matrix, receive the data from the master node, multiply it with the key node, stores the result into a variable and send back the resultant data to the master node. To utilize multiple nodes simultaneously, a PBS shell script is used. The cryptanalysis of various lengths of ciphers and their respective time to decipher all possible plaintext using MPI model is shown in Table 3.

3.1.2 CUDA + MPI model on SIGABA, AES, ZUC & SNOW ciphers

This is the first HPC based attack with CUDA+MPI on SIGABA cipher. The Key space of a physical SIGABA machine with a fixed mechanical rotor is no more than $2^{95.6}$. For a known cipher rotor initial setting and a fixed order of cipher & control rotor, the key space will reduce to $2^{50.3}$. For cryptanalysis, SIGABA ciphertext of 250-character is used. This is first comparative strength analysis of 4G & 5G Mobile communication Ciphers using CUDA+MPI model on HPC. AES using a 128-bit key that needs ten rounds, with each round requiring 16 bytes of round key data, a 128-bit key version of SNOW & a 128-bit key version of ZUC are the mobile communication (4G & 5G) ciphers used in this analysis. A brute-force attack is performed on these ciphers by setting the key in such a way that, it is found after & close to 2^{50} iterations. To utilize multiple nodes simultaneously, a PBS shell script is used. The time taken for the attack on all 4 ciphers is shown in Table 4.

Table 4. Analysis of SIGABA, AES, ZUC & SNOW ciphers using CUDA + MPI

Cipher	Key strength(in power of 2 i.e., 2^X)	Exhausted key space	No of compute nodes used	Resources used per node	Total no of resources used	Time taken (in days.hours)
SIGABA	95.6	$\approx 2^{50}$	10	1 CPU &2 GPU	10 CPU &20 GPU	3.6
AES	128					5.4
SNOW	128					12.1
ZUC	128					15.7

Considering the computational power of HPC machine used in this work (100TFLOPS) and the current HPC machines across the globe (Current world's fastest HPC-FRONTIER, Oak Ridge National Laboratory, US, with an approximate speed of 1 EFLOPS), it is safe to say that mobile communication ciphers are not at risk from exhaustive key search attacks.

3.2 Machine Learning-based Cryptanalysis

The following Classification and Clustering techniques have not applied on the following ciphers before this work.

3.2.1 Clustering of machine ciphers data

Clustering of image data for various distance measures was previously performed by Ajeet Singh et. al.[5]. Figures 1 and 2 & Figure 3. show the results of the clustering models, K Means and Hierarchical clustering after extracting various features from the data generated from image ciphertext of 3 machine Ciphers i.e., Enigma, Fialka, KL-7. For finding optimal hyper parameter, silhouette score is used for hierarchical clustering and Elbow method is used for K-means clustering. The features used in this procedure include n-gram frequency (where, n = {1,2,3}) and TF-IDF.

3.2.2 Classification of machine ciphers data

Ju-Won Park et al., have used various features related to I/O operations like totalFile, totalOpenReq, seqWritePct, runtime, etc., using five different feature selection methods[4]. Table 5 shows, results of the various classification models after extracting features from the data generated from image ciphertext of 3 machine ciphers. The optimal hyper parameter is found in each case after experimenting with several values. The features used in this procedure are n-gram frequency (where, n = {1,2,3}) and TF-IDF. Out of 5 classification algorithms used, Logistic Regression offered highest accuracy upto 93.9% while Naïve Bayes offered lowest upto 62.4%.

3.2.3 Classification of mobile communication ciphers data

S. Ramraj & G. Usha has tried classification of Watsapp media (image/text) content using various machine learning models[7]. V. Nalla et al., previous work has 10000 ciphertext of all three ciphers as input data but here 20000 ciphertext of the three ciphers are generated as data[16]. The results of the various classification models after extracting features from the data generated from

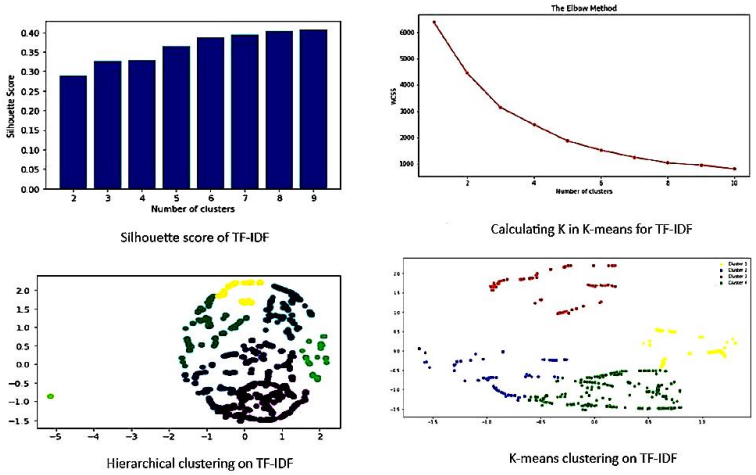

Fig 1. Results of the clustering models after extracting TF-IDF feature from machine cipher data

20000 ciphertext of each of the three mobile communication ciphers – AES, SNOW and ZUC are shown here. The optimal hyper parameter is found in each case after experimenting with several values. Table 6 shows performance metrics of Entropy-based analysis in which, out of 4 classification algorithms used, KNN offered highest accuracy of 78.3% while SVM offered lowest of 66.7%. Table 7 shows performance metrics of Long Repeat, Index of Coincidence, Number of Unique Characters and Di-graphic IoC in which, out of 4 classification algorithms used, KNN offered highest accuracy of 79.2% while SVM offered lowest of 66.7%. Table 8 shows the performance metrics of combination of both above procedures i.e., Entropy, Long Repeat, Index of Coincidence, Number of Unique Characters and Di-graphic IoC in which, out of 4 classification algorithms used, KNN offered highest accuracy of 75.8% while SVM offered lowest of 66.6%.

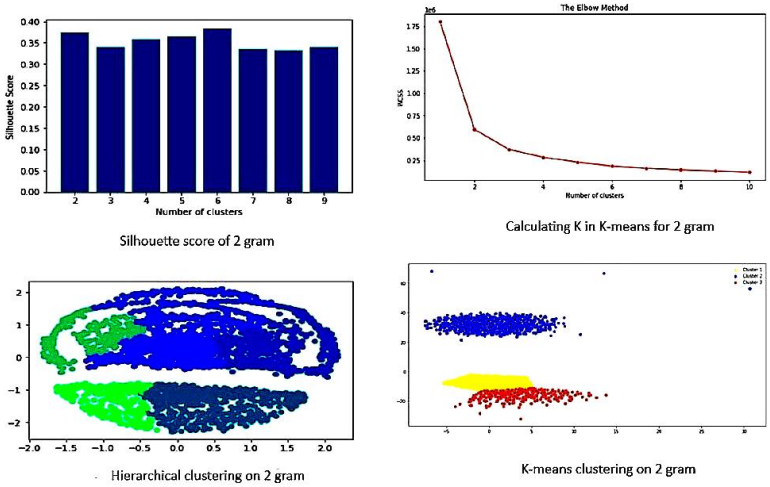

Fig 2. Results of the various clustering models after extracting 2-gram frequency feature from machine cipher data

As an extension to this work, these algorithms resistance can be analyzed on upscaled HPCs with higher performance[17]. More features can be added to classification algorithms to improve accuracy. The complex neural deep learning model and GAN model may be explored to improve the accuracy[18]. One can also investigate ML's capacity to generate computationally effective stand-in models of real-world applications, obviating the need for more costly simulation methods altogether[19]. These

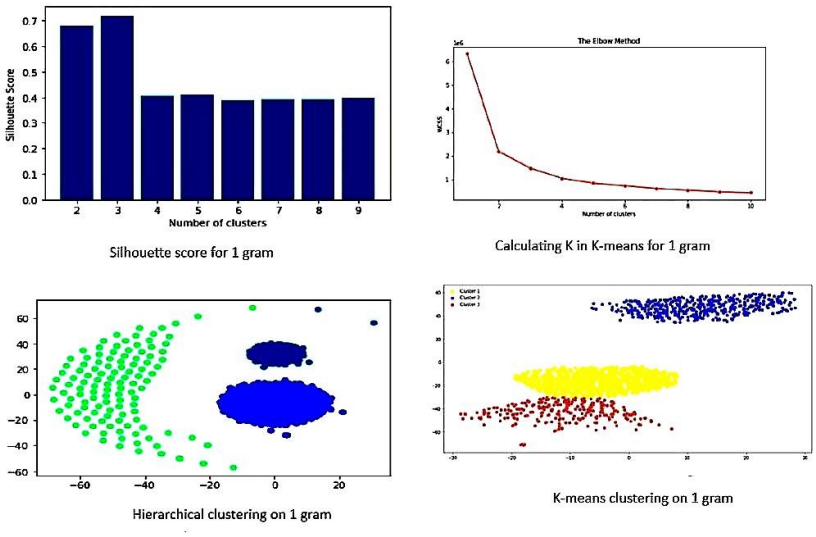

Fig 3. Results of the various clustering models after extracting 1-gram frequency feature from machine cipher data

Table 5. Various classification models after extracting features from machine cipher data

Classification Model	Text Feature	Hyper parameter Depth (K)	Training accuracy	Testing accuracy
KNN	1 gram	9	77.2	74.8
KNN	2 gram	7	80.3	80.6
KNN	3 gram	5	80.5	80.9
KNN	TF-IDF	4	88.1	87.2
Naïve Bayes	1 gram	60	77.4	76.9
Naïve Bayes	2 gram	50	80.5	81.2
Naïve Bayes	3 gram	20	62.5	62.4
Naïve Bayes	TF-IDF	0	80.3	80.1
Logistic Regression	1 gram	2	100	93.8
Logistic Regression	2 gram	1	100	93.9
Logistic Regression	TF-IDF	1	94.4	88.3
Decision Tree	1 gram	5	97.0	87.7
Decision Tree	2 gram	8	100	88.8
Decision Tree	3 gram	3	95.1	80.4
Decision Tree	TF-IDF	4	95.2	93.6
SVM	1 gram	5	87.1	87.2
SVM	2 gram	4	87.2	88.2
SVM	3 gram	3	89.0	87.6
SVM	TF-IDF	19	74.8	69.0

Table 6. Performance metrics of Entropy-based analysis

Classification Technique	Training Accuracy	Testing Accuracy	Precision			Recall			F1-score		
			Label-1	Label-2	Label-3	Label-1	Label-2	Label-3	Label-1	Label-2	Label-3
KNN	78.3	67.6	1.00	0.51	0.52	1.00	0.52	0.51	1.00	0.52	0.51
Decision Tree	70.5	66.6	1.00	0.50	0.51	1.00	0.71	0.30	1.00	0.59	0.38
Random Forest	70.4	66.5	1.00	0.50	0.50	1.00	0.52	0.48	1.00	0.51	0.49
SVM	66.7	66.6	1.00	0.50	0.51	1.00	0.90	0.10	1.00	0.64	0.17

Table 7. Performance metrics of Features only

Classification Technique	Training Accuracy	Testing Accuracy	Precision			Recall			F1-score		
			Label-1	Label-2	Label-3	Label-1	Label-2	Label-3	Label-1	Label-2	Label-3
KNN	79.2	66.3	1.00	0.49	0.50	1.00	0.51	0.48	1.00	0.50	0.49
Decision Tree	70.67	67.4	1.00	0.48	0.48	1.00	0.47	0.48	1.00	0.48	0.48
Random Forest	70.67	67.0	1.00	0.50	0.51	1.00	0.58	0.43	1.00	0.54	0.47
SVM	66.7	66.7	1.00	0.00	0.50	1.00	0.00	1.00	1.00	0.00	0.67

Table 8. Performance metrics in combination of Entropy and features

Classification Technique	Training Accuracy	Testing Accuracy	Precision			Recall			F1-score		
			Label-1	Label-	Label-3	Label-1	Label-2	Label-3	Label-1	Label-2	Label-3
KNN	75.8	66.8	1.00	0.50	0.50	1.00	0.50	0.51	1.00	0.50	0.51
Decision Tree	86.9	66.8	1.00	0.48	0.49	1.00	0.47	0.49	1.00	0.48	0.49
Random Forest	86.8	66.8	1.00	0.50	0.50	1.00	0.50	0.51	1.00	0.50	0.51
SVM	66.6	66.4	1.00	0.41	0.50	1.00	0.02	0.97	1.00	0.04	0.66

ML techniques may be tried for classification of newly designed resource-constraint ciphers like Espresso, m-Crypton, grain ciphers[20,21].

4 Conclusion

This study explored the strength of cryptographic applications using HPC techniques, focusing on the MPI model and the CUDA + MPI model. The MPI model is applied exclusively to the Hill cipher (for key sizes of 2x2, 3x3 & 4x4). The CUDA + MPI model is implemented for the SIGABA cipher & Mobile communication ciphers with 60K plaintext data which was not done before this work. Considering the power of current HPC machines across the globe, it is safe to say that mobile communication ciphers have no danger from exhaustive key search attacks. The investigation of the various classification and clustering models on popular machine ciphers is done using encrypted image data collected from online with test accuracy up to 93.9%. This work shows the effect of ML techniques in cryptanalysis even though the analysis on machine ciphers which are seldom used in current era.

Finally, the study explored machine learning methodologies for the classification of 4G and 5G mobile communication ciphers solely based on ciphertext analysis with test accuracy up to 79.2%. More data from the algorithms (> 1 million per algorithm) might improve the accuracy of classification. Ananth R & Ramaiah N have analyzed various stream ciphers on multiple parameters which may be used for classification and/or clustering them, as a future work[2].

Acknowledgements

This paper has been presented in "International Conference on Applied Mathematics and Advanced Data Analytics for Industry 5.0 (ICAMADA-2024)" held during 25th to 27th April 2024 at VNR VJIET, in collaboration with APTSMS and CCOE. The authors express their sincere gratitude to the guest editors for their tireless efforts and dedication by means of comments, editing and overseeing the manuscripts to bring in this final form.

The APC is deferred partially by Indian Society for Education and Environment.

References

1) Gill SS, Wu H, Patros P, Ottaviani C, Arora P, Pujol VC, et al. Modern computing: Vision and challenges. *Telematics and Informatics Reports*. 2024;13. Available from: https://dx.doi.org/10.1016/j.teler.2024.100116.
2) Ananth R, Ramaiah NS. An exhaustive review of the stream ciphers and their performance analysis. *International Journal of Reconfigurable and Embedded Systems (IJRES)*. 2024;13(2). Available from: https://dx.doi.org/10.11591/ijres.v13.i2.pp360-371.
3) Park JW, Huang X, Lee JK, Hong T. I/O-signature-based feature analysis and classification of high-performance computing applications. *Cluster Computing*. 2024;27(3):3219–3231. Available from: https://dx.doi.org/10.1007/s10586-023-04139-y.

4) Bang J, Kim C, Wu K, Sim A, Byna S, Kim S, et al. HPC Workload Characterization Using Feature Selection and Clustering. *Proceedings of the 3rd International Workshop on Systems and Network Telemetry and Analytics*. 2020;p. 33–40. Available from: https://dl.acm.org/doi/10.1145/3391812.3396270.
5) Singh A, Sivangi KB, Tentu AN. Machine Learning and Cryptanalysis: An In-Depth Exploration of Current Practices and Future Potential. *Journal of Computing Theories and Applications*. 2024;1(3):257–272. Available from: https://dx.doi.org/10.62411/jcta.9851.
6) Schleider L, Pasiliao EL, Qiang Z, Zheng QP. A study of feature representation via neural network feature extraction and weighted distance for clustering. *Journal of Combinatorial Optimization*. 2022;44(4):3083–3105. Available from: https://dx.doi.org/10.1007/s10878-022-00849-y.
7) Ramraj S, Usha G. Hybrid feature learning framework for the classification of encrypted network traffic. *Connection Science*. 2023;35(1). Available from: https://dx.doi.org/10.1080/09540091.2023.2197172.
8) Kavitha T, Rajitha O, Thejaswi K, Muppalaneni NB. Classification of Encryption Algorithms Based on Ciphertext Using Pattern Recognition Techniques. In: Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018). Springer International Publishing. 2020;p. 540–545. Available from: https://www.springerprofessional.de/en/classification-of-encryption-algorithms-based-on-ciphertext-usin/17027850#:~:text=The%20classification%20techniques%20used%20are,algorithm%20for%20the%20given%20ciphertext.
9) Mukherjee CS, Roy D, Maitra S. Design Specification of ZUC Stream Cipher. In: SpringerBriefs on Cyber Security Systems and Networks. Springer Singapore. 2021;p. 43–62. Available from: https://link.springer.com/chapter/10.1007/978-981-33-4882-0_3#:~:text=In%20the%20key%2DIV%20initialization,initialized%20to%20some%20padding%20bits.
10) Lasry G. Cracking SIGABA in less than 24 hours on a consumer PC. *Cryptologia*. 2023;47(1):1–37. Available from: https://doi.org/10.1080/01611194.2021.1989522.
11) Saputra DM, Saputra D, Oswari LD. Effect of Distance Metrics in Determining K-Value in K-Means Clustering Using Elbow and Silhouette Method. *Proceedings of the Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*. 2020. Available from: https://www.atlantis-press.com/proceedings/siconian-19/125939938.
12) Ashari IF, Nugroho ED, Baraku R, Yanda IN, Liwardana R. Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta. *Journal of Applied Informatics and Computing*. 2023;7(1):89–97. Available from: https://dx.doi.org/10.30871/jaic.v7i1.4947.
13) Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*. 2021;2(160). Available from: https://link.springer.com/article/10.1007/s42979-021-00592-x.
14) Khan A, Qureshi M, Daniyal M, Tawiah K. A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction. *Health & Social Care in the Community*. 2023;p. 1–10. Available from: https://dx.doi.org/10.1155/2023/1406060.
15) Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Frontiers in Bioinformatics*. 2022;2. Available from: https://dx.doi.org/10.3389/fbinf.2022.927312.
16) Nalla V, Pooja C, Padmavathi G, Kameswari US. Classification of Ciphers using Only Cipher Text for Ciphers used in 4G & 5G Mobile Communication. In: 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI). IEEE. 2024;p. 720–728. Available from: https://ieeexplore.ieee.org/document/10493871.
17) Azad M, Iqbal N, Hassan F, Roy P. An empirical study of high performance computing (HPC) performance bugs. In: 2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR). 2023;p. 194–206. Available from: https://ieeexplore.ieee.org/document/10174003.
18) Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. *Communications of the ACM*. 2020;63(11):139–144. Available from: https://dx.doi.org/10.1145/3422622.
19) Frank M, Drikakis D, Charissis V. Machine-Learning Methods for Computational Science and Engineering. *Computation*. 2020;8(1). Available from: https://dx.doi.org/10.3390/computation8010015.
20) Kamalanathan C, Balamurugan J, Sharma N, Reddy AB, Selvan RS. Development of Lightweight and Cheaper 5G Mobile Communication System to Analyze the Performance of Espresso Ciphers and Grain Family. In: Communications in Computer and Information Science. Springer Nature Switzerland. 2025;p. 140–153. Available from: https://link.springer.com/chapter/10.1007/978-3-031-73494-6_10.
21) Singh P, Prasad SVS, Upadhyay S, Singh R. Performance-efficient flexible architecture of m–Crypton cipher for resource-constrained applications. *Automatika*. 2024;65(4):1447–1457. Available from: https://dx.doi.org/10.1080/00051144.2024.2395617.