Robust Botnet Dga Detection: Blending Xai And Osint For Cyber Threat Intelligence Sharing

Mr. M. Sudhakar[1*], Mrs. Srilatha Puli[2], Rachana Reddy Sunki[3], T. Sri Anjali[4], P. Vamshi Vardhan[5], T. Rekha[6],

[1*]Assistant Professor, Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email: ios.yadav143@gmail.com
[2]Assistant Professor, Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email: srilatha.puli@sreyas.ac.in
[3]Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email: rachanareddysunki@gmail.com
[4]Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email srianjali9080@gmail.com
[5]Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email: pvamshivardhan789@gmail.com
[6]Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email: tadipatrirekha1998@gmail.com

***Corresponding Author:** Mr. M. SUDHAKAR
*Assistant Professor, Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India., Email: ios.yadav143@gmail.com

Abstract
We investigated 12 years DNS query logs of our campus network and identified phenomena of malicious botnet domain generation algorithm (DGA) traffic. DGA-based botnets are difficult to detect using cyber threat intelligence (CTI) systems based on blocklists. Artificial intelligence (AI)/machine learning (ML)-based CTI systems are required. This study (1) proposed a model to detect DGA-based traffic based on statistical features with datasets comprising 55 DGA families, (2) discussed how CTI can be expanded with computable CTI paradigm, and (3) described how to improve the explain ability of the model outputs by blending explainable AI (XAI) and open-source intelligence (OSINT) for trust problems, an antidote for skepticism to the shared models and preventing automation bias. We define the XAI-OSINT blending as aggregations of OSINT for AI/ML model outcome validation. Experimental results show the effectiveness of our models (96.3% accuracy). Our random forest model provides better robustness against three state of-the-art DGA adversarial attacks (Char Bot, Deep DGA, Mask DGA) compared with character-based deep learning models (Endgame, CMU, NYU, MIT). We demonstrate the sharing mechanism and confirm that the XAI-OSINT blending improves trust for CTI sharing as evidence to validate our proposed computable CTI paradigm to assist security analysts in security operations centers using an automated, explainable OSINT approach (for second opinion). Therefore, the computable CTI reduces manual intervention in critical cybersecurity decision-making.

1. INTRODUCTION

Botnets are powerful tools used by cybercriminals to carry out various malicious activities, such as distributed denial-of-service (DDoS) attacks, spam campaigns, and data theft. Detecting and mitigating botnets is a crucial task for cybersecurity professionals to protect networks and systems from their harmful effects. One technique commonly employed by botnets is the Domain Generation Algorithm (DGA), which generates a large number of domain names to establish command and control (C&C) communication channels**.**

To effectively detect and counter botnets that utilize DGAs, a robust approach is required, blending cutting-edge technologies such as Explainable Artificial Intelligence (XAI) and Open-Source Intelligence (OSINT) with Cyber Threat Intelligence (CTI). This integration enables cybersecurity experts to leverage advanced analytics, machine learning, and real-time information to enhance their botnet detection capabilities**.** Explainable Artificial Intelligence (XAI) plays a crucial role in understanding the inner workings of machine learning models. By using XAI techniques, cybersecurity professionals can interpret the decisions and predictions made by AI systems, making them more transparent and trustworthy.

Open-Source Intelligence (OSINT) involves collecting and analyzing publicly available data from various sources, such as social media, websites, and forums. OSINT provides valuable information about known botnet activities, indicators of compromise (IoCs), and the tactics, techniques, and procedures (TTPs) employed by cybercriminals. Cyber Threat Intelligence (CTI) refers to the knowledge and insights gained from analyzing data related to cyber threats and adversaries.

It includes information about attacker profiles, their motivations, and their methods. By leveraging CTI, cybersecurity professionals can proactively anticipate and respond to botnet activities, enhancing their detection and mitigation capabilities. Blending XAI, OSINT, and CTI enables a comprehensive and proactive approach to botnet detection. By combining the interpretability of XAI techniques, the real-time information from OSINT sources, and the strategic insights from CTI, cybersecurity professionals can build more effective and resilient defense mechanisms against botnet threats.

2.LITERATURE REVIEW

A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing:

The Internet threat landscape is fundamentally changing. A major shift away from hobby hacking toward well-organized cyber crime can be observed. These attacks are typically carried out for commercial reasons in a sophisticated and targeted manner, and specifically in a way to circumvent common security measures. Additionally, networks have grown to a scale and complexity, and have reached a degree of interconnectedness, that their protection can often only be guaranteed and financed as shared efforts. Consequently, new paradigms are required for detecting contemporary attacks and mitigating their effects. Today, many attack detection tasks are performed within individual organizations, and there is little cross-organizational information sharing. However, information sharing is a crucial step to acquiring a thorough understanding of large-scale cyber-attack situations, and is therefore seen as one of the key concepts to protect future networks.

A historical perspective of explainable artificial intelligence:

Explainability in Artificial Intelligence (AI) has been revived as a topic of active research by the need of conveying safety and trust to users in the "how" and "why" of automated decision-making in different applications such as autonomous driving, medical diagnosis, or banking and finance. While explainability in AI has recently received significant attention, the origins of this line of work go back several decades to when AI systems were mainly developed as (knowledge-based) expert systems. Since then, the definition, understanding, and implementation of explainability have been picked up in several lines of research work, namely, expert systems, machine learning, recommender systems, and in approaches to neural-symbolic learning and reasoning, mostly happening during different periods of AI history.

A survey of explainable AI terminology:

The field of Explainable Artificial Intelligence attempts to solve the problem of algorithmic opacity. Many terms and notions have been introduced recently to define Explainable AI, however, these terms seem to be used interchangeably, which is leading to confusion in this rapidly expanding field. As a solution to overcome this problem, we present an analysis of the existing research literature and examine how key terms, such as transparency, intelligibility, interpretability, and explainability are referred to and in what context. This paper, thus, moves towards a standard terminology for Explainable AI. Explainable AI, black-box, NLG, Theoretical Issues, Transparency, Intelligibility, Interpretability, Explainability, XAI.

A bibliometric analysis of the explainable artificial intelligence research field:

This paper presents the results of a bibliometric study of the recent research on explainable Artificial Intelligence (XAI) systems. We took a global look at the contributions of scholars in XAI as well as in the subfields of AI that are mostly involved in the development of XAI systems. It is worthy to remark that we found out that about one third of contributions in XAI come from the fuzzy logic community. Accordingly, we went in depth with the actual connections of fuzzy logic contributions with AI to promote and improve XAI systems in the broad sense. Finally, we outlined new research directions aimed at strengthening the integration of different fields of AI, including fuzzy logic, toward the common objective of making AI accessible to people.

Issues and challenges in DNS based botnet detection: A survey:

Cybercrimes are evolving on a regular basis and as such these crimes are becoming a greater threat day by day. Earlier these threats were very general and unorganized. In the last decade, these attacks have become highly sophisticated in nature. This higher level of coordination is possible mainly due to botnets, which are clusters of infected hosts controlled remotely by an attacker (botmaster). The number of infected machines is continuously rising, thereby resulting in botnets with over a million infected machines. This powerful capability gives the botmaster a lethal weapon to launch various security attacks.

3. METHODOLOGY

Botnet Domain Generation Algorithm (DGA) is used to generate malicious domain which can be used to steal information from normal user system. DGA will generate malicious URL and this URL will allure normal user by sending messages like lottery win or any other fake message. Whenever user click on such malicious domain then that domain will install malicious program in user system and then start stealing secret data or perform any other malicious activity. DAG can be detected using machine learning or deep learning algorithms and it's difficult to detect with Cyber Threat Intelligence (CTI).

The flaws in the current system

it's difficult to detect with Cyber Threat Intelligence (CTI)

Due to the characteristics of the algorithmically generated domain in the DGA, depending only on a domain blocklist might be insufficient.

To detect above malicious in this paper employing Machine Learning algorithms such as Random Forest, Naïve Bayes, Logistic Regression, Extra Tree and Ensemble algorithm. In all algorithms Random Forest is giving best accuracy and then employing XAI (explainable ML algorithm) technique which will explain about the performance of the machine learning model. XAI will display summary of all prediction and what percentage of impact make the ML model to predict such class label.

1. Experimental results show the effectiveness of our models (96.3% accuracy).
2. Our random forest model provides better robustness

4.IMPLEMENTATION

To train and test ML model author using online social network (OSINT) trust model dataset and then select 6 important features using Shanon Entropy calculation and those selected Features.

We have used BOTNET real dataset generated from their college campus and this dataset can be downloaded from IEEE data port. In below screen showing dataset details

In above dataset screen first row contains dataset column names and remaining rows contains dataset values and in last column we have class label called DA which contains values as 0 or 1 where 0 means normal DOMAIN and 1 means malicious BOTNET domain.

Extension Concept: in propose paper author has used all traditional algorithms to detect DGA so as extension we have employed advance machine learning algorithm called XGBOOST which will use multiple estimators to train and test class label so its accuracy will be higher compare to other algorithms

To implement this project we have designed following modules

1) Upload Botnet DGA Dataset: using this module we will upload dataset to application and then find and plot different class labels found in dataset
2) Preprocess Dataset: using this module we will read dataset and then remove missing values and then shuffle, normalized and split dataset into train and test where application using 80% dataset for training and 20 for testing
3) Run Random Forest: 80% processed dataset will be input to Random Forest algorithm to train a model and this model will be applied on 20% test data to calculate random forest prediction accuracy
4) Run Logistic Regression: 80% processed dataset will be input to Logistic Regression algorithm to train a model and this model will be applied on 20% test data to calculate LR prediction accuracy
5) Run Naive Bayes: 80% processed dataset will be input to Naïve Bayes algorithm to train a model and this model will be applied on 20% test data to calculate NB prediction accuracy
6) Run Extra Tree: 80% processed dataset will be input to Extra Tree algorithm to train a model and this model will be applied on 20% test data to calculate ET prediction accuracy
7) Run Ensemble Algorithm: 80% processed dataset will be input to Ensemble algorithm to train a model and this model will be applied on 20% test data to calculate Ensemble prediction accuracy
8) Run Extension XGBoost: 80% processed dataset will be input to XGBOOST algorithm to train a model and this model will be applied on 20% test data to calculate XGB prediction accuracy
9) Shapely XAI Graph: using this module we will plot XAI graph which will explain about algorithm prediction.
10) Predict DGA Botnet from Test Data: using this module we will upload test data and then ML algorithm will predict weather test data is Normal Domain or Malicious DGA.

5. EXPERIMENTAL RESULTS

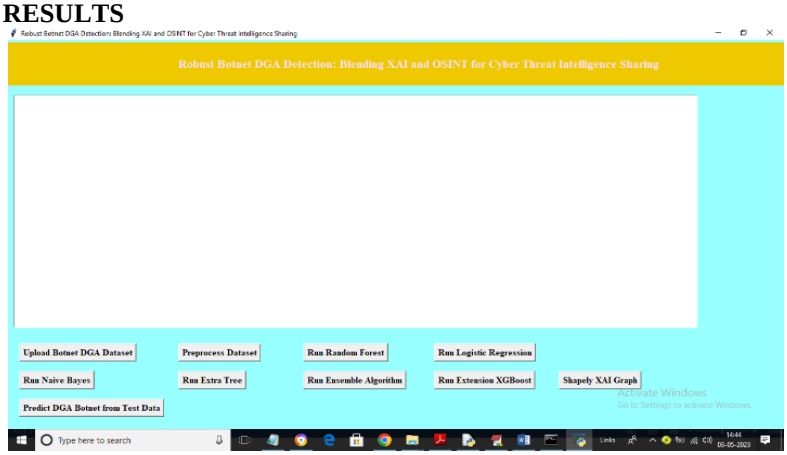

Fig 1 : output

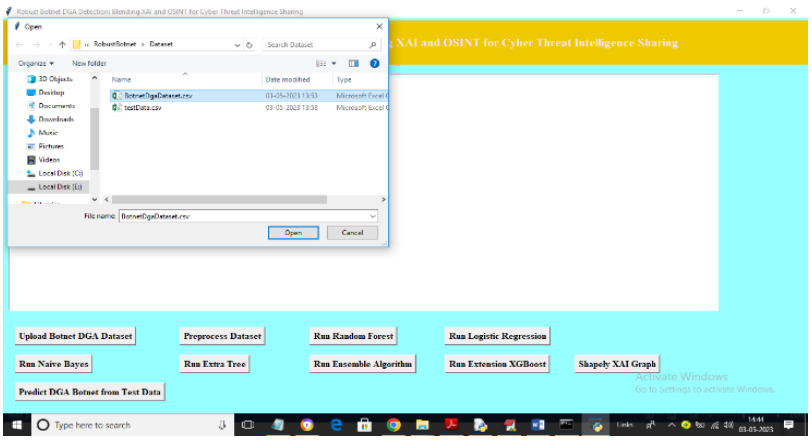

Fig 2: upload botnet DGA dataset

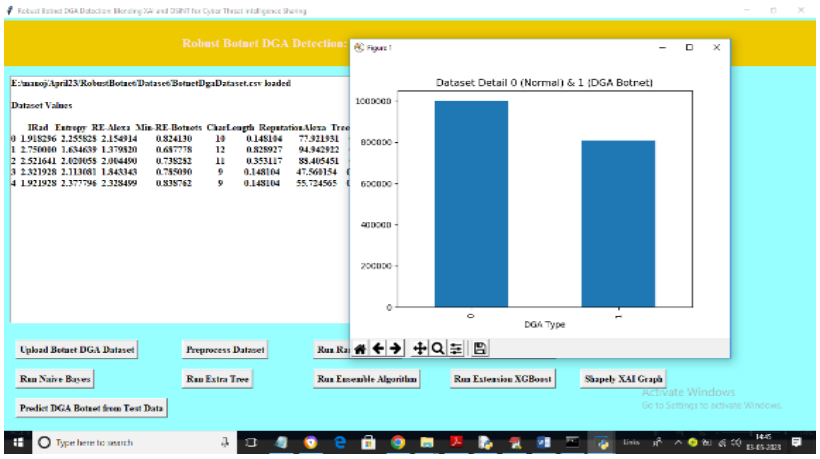

Fig 3 : DGA Type

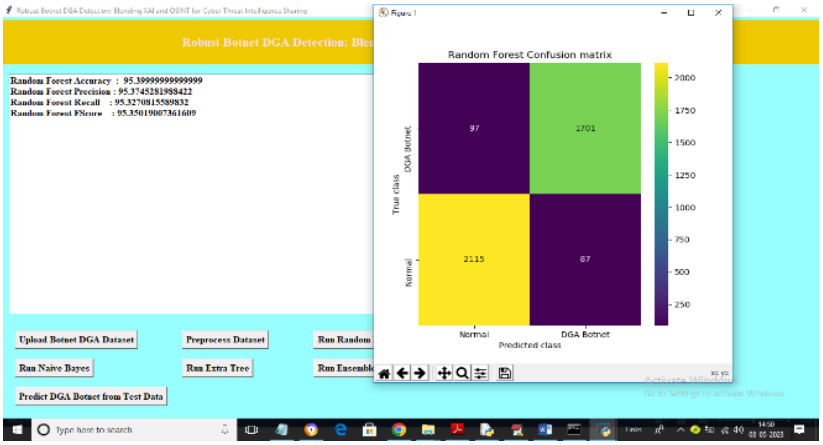

Fig 4: Random forest

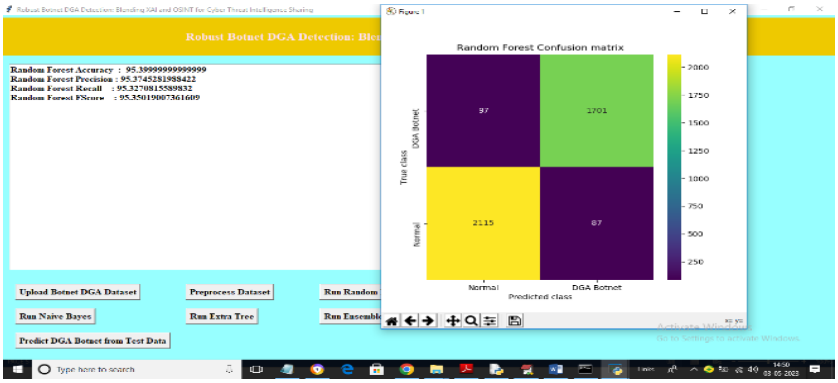

Fig 5: logistic regression

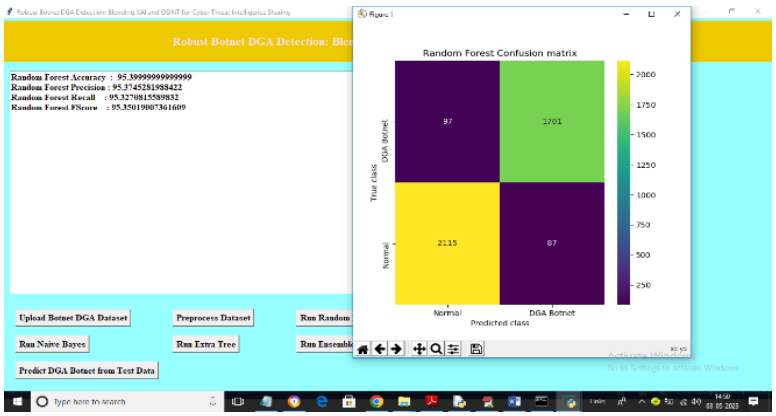

Fig 6: Naïve Bayes

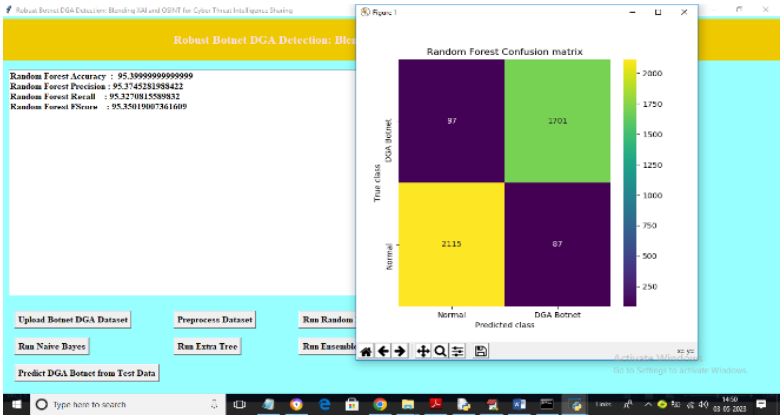

Fig 7 : extra tree

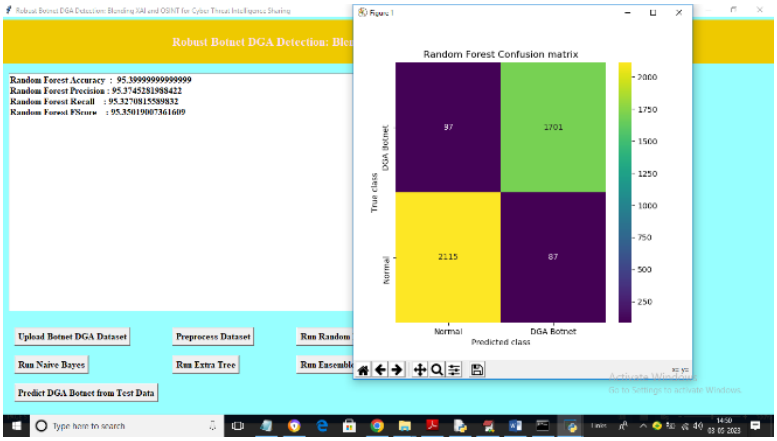

Fig 8: XGBoost

Fig 9: XAI graph

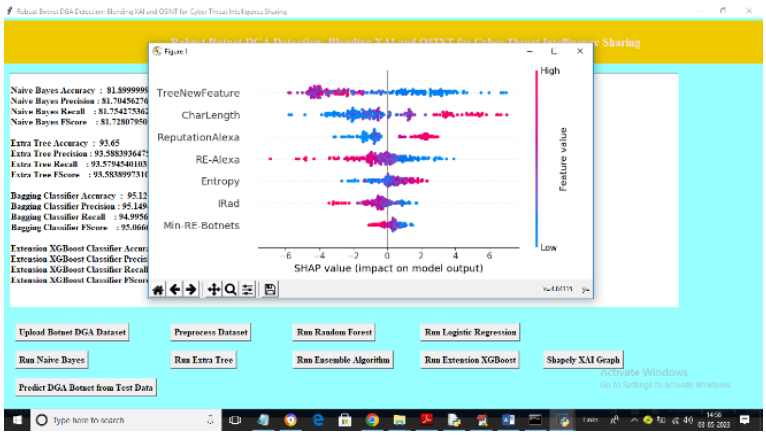

Fig 10: XAI graph

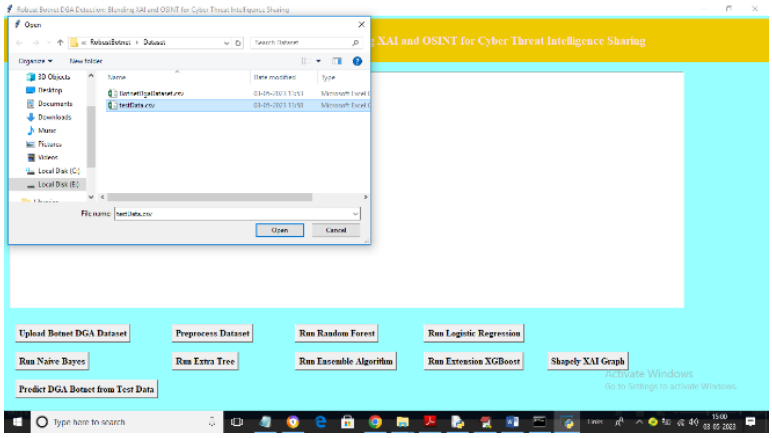

Fig 11: Test data

Fig 12: prediction of Botnet

6.CONCLUSION

First, we showcase a novel model for botnet DGA detection. Our random forest model achieved 96.3% accuracy and outperformed the previous work Our model is also more robust against three state-of-the-art DGA adversarial attacks than the previous works Second, we highlight the practicality of blending XAI and OSINT to deliver better AI explainability through second opinion approaches, thus mimicking the second opinion phenomena in hospital/medical situations to confirm the results/findings. We advocate the XAI and OSINT as an antidote for skepticism toward the model's output, which might contribute to the CTI system's trust and prevent automation bias when users have too much trust in the CTI system's output. Blending XAI and OSINT also has a potential for solving the false-flag problems. Third, we underline the case study of botnet DGA detection with XAI and OSINT blend as evidence to validate our proposed computable CTI paradigm. Improving trust might result in a paradigm-shift phenomenon. Cybersecurity communities will leave the traditional CTI-sharing paradigm (sharing only threat indicators, such as threat domain names), and communities will start to share AI/ML models for CTI systems. With the emergence of the computable CTI-sharing paradigm, additional collaboration among cybersecurity communities will occur to develop advanced AI/ML-based CTI systems. For instance, using transfer-learning techniques to develop new AI/ML for new cybersecurity tasks/problems utilizing the shared models. The limitations of our DGA detection model are the time complexity when calculating the features and the limited robustness against MaskDGA attacks Future improvement should focus on crafting better features and adversarial defense

strategies. Moving target defense (MTD) [63] can potentially raise the model's robustness by combining various models to work together.

7.REFERENCES

1. F. Skopik, G. Settanni, and R. Fiedler, ''A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing,'' Comput. Secur., vol. 60, pp. 154–176, Jul. 2016, doi: 10.1016/j.cose.2016.04.003.
2. R. Confalonieri, L. Coba, B. Wagner, and T. R. Besold, ''A historical perspective of explainable artificial intelligence,'' WIREs Data Mining Knowl. Discovery, vol. 11, no. 1, p. e1391, Jan. 2021, doi: 10.1002/widm.1391.
3. M.-A. Clinciu and H. Hastie, ''A survey of explainable AI terminology,'' in Proc. 1st Workshop Interact. Natural Lang. Technol. Explainable Artif. Intell. (NLXAI), 2019, pp. 8–13, doi: 10.18653/v1/W19-8403.
4. J. M. Alonso, C. Castiello, and C. Mencar, ''A bibliometric analysis of the explainable artificial intelligence research field,'' in Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations. Cham, Switzerland: Springer, 2018, pp. 3–15, doi: 10.1007/978-3-319-91473-2_1.
5. M. Singh, M. Singh, and S. Kaur, ''Issues and challenges in DNS based botnet detection: A survey,'' Comput. Secur., vol. 86, pp. 28–52, Sep. 2019, doi: 10.1016/j.cose.2019.05.019.
6. T. S. Wang, H.-T. Lin, W.-T. Cheng, and C.-Y. Chen, ''DBod: Clustering and detecting DGA-based botnets using DNS traffic analysis,'' Comput. Secur., vol. 64, pp. 1–15, Jan. 2017, doi: 10.1016/j.cose.2016.10.001.
7. X. D. Hoang and X. H. Vu, ''An improved model for detecting DGA botnets using random forest algorithm,'' Inf. Secur. J., Global Perspective, pp. 1–10, Jun. 2021, doi: 10.1080/19393555.2021.1934198.
8. B. Yu, J. Pan, J. Hu, A. Nascimento, and M. De Cock, ''Character level based detection of DGA domain names,'' in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2018, pp. 1–8, doi: 10.1109/IJCNN.2018.8489147.
9. L. Sidi, A. Nadler, and A. Shabtai, ''MaskDGA: An evasion attack against DGA classifiers and adversarial defenses,'' IEEE Access, vol. 8, pp. 161580–161592, 2020, doi: 10.1109/ACCESS.2020.3020964.
10. J. Peck, C. Nie, R. Sivaguru, C. Grumer, F. Olumofin, B. Yu, A. Nascimento, and M. De Cock, ''CharBot: A simple and effective method for evading DGA classifiers,'' IEEE Access, vol. 7, pp. 91759–91771, 2019, doi: 10.1109/ACCESS.2019.2927075.
11. H. S. Anderson, J. Woodbridge, and B. Filar, ''DeepDGA: Adversariallytuned domain generation and detection,'' in Proc. ACM Workshop Artif. Intell. Secur., Vienna, Austria, Oct. 2016, pp. 13–21, doi: 10.1145/2996758.2996767.
12. T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah, ''Cyber threat intelligence sharing: Survey and research directions,'' Comput. Secur., vol. 87, Nov. 2019, Art. no. 101589, doi: 10.1016/j.cose.2019.101589.
13. W. Tounsi and H. Rais, ''A survey on technical threat intelligence in the age of sophisticated cyber attacks,'' Comput. Secur., vol. 72, pp. 212–233, Jan. 2018, doi: 10.1016/j.cose.2017.09.001.
14. P. Pawlinski, P. Jaroszewski, P. Kijewski, L. Siewierski, P. Jacewicz, P. Zielony, and R. Zuber, ''Actionable information for security incident response,'' in Proc. Eur. Union Agency Netw. Inf. Secur., Heraklion, Greece, 2014, pp. 1–68.
15. M. Sudhakar Published a paper entitled "EMOTION BASED MUSIC PLAYERA" In (Volume 21, May- 2022 Issue05, YMER).
16. M. Sudhakar Published a paper entitled "CASHLESS SOCIETY: MANAGING PRAVACY AND SECURITYA" In (Volume 02, July- 2022 Issue, YMER).
17. M. Sudhakar published a paper entitled "DETECTING A POTHOLE USING DEEP CNN FOR AN ADAPTIVE SHOCK OBSERVING IN A VECHILE DRIVINGE" In (Volume 20, June- 2022 Issue06, NERO QUANTOLOGY SCOPUS).
18. M. Sudhakar published a paper entitled "DETECTION OF NEUROLOGICAL DISORDER" In (Volume 13, Jan-2023 Issue01, IJARST).
19. M. Sudhakar Published a paper entitled "An Ensemble Algorithm for Crop Yield Prediction in Agriculture Sector" In (Volume 15, Jan- 2023 Issue01,JICR).
20. M. Sudhakar Published a paper entitled "COVID-19 Epidemic Analysis using Machine Learing and Deep Learning Algorithm" In (Volume 05, Issue: 08 AUG 2021, IJSREM).
21. M. Sudhakar Published a paper entitled "A SURVEY ON IMPROVED PERFORMANCE FOR KEYWORD QUERY ROUTING" In (Volume 02, July- 2015 Issue, JREECSM).
22. M. Sudhakar Published a paper entitled "A Method for Forecasting Heart Disease using Effective Machine Learning Process" at the (ICRCSIT-20) held on June 17th and 18th 28, 2020.