

Topological dynamics algorithms for enhanced cybersecurity protocols

Reena Mahapatra Lenka [‡]

*Symbiosis Institute of Management Studies
Symbiosis International (Deemed) University
Pune 411020
Maharashtra
India*

Sudhakar Tiwari [†]

*Schaumburg
Illinois 60195
United States*

Sandeep Dommari [§]

*Prosper
Texas-75078
United States*

Varsha Kiran Bhosale ^{*}

*Department of Computer Science and Engineering
Arvind Gavali College of Engineering
Satara
Maharashtra
India*

Pranav Dass [®]

*Department of Computer Science
Shyam Lal College
University of Delhi
Delhi 110032
India*

[‡] E-mail: reena.lenka@sims.edu

[†] E-mail: sudhakar.tiwari@ieee.org

[§] E-mail: sandeep.dommari@ieee.org

^{*} E-mail: vkbhosale21@gmail.com (Corresponding Author)

[®] E-mail: pdasscs@shyamlal.du.ac.in

Nikita Chavan [#]

*Department of Electronics and Telecommunication
Dr. D. Y. Patil Institute of Technology
Pimpri
Pune
Maharashtra
India*

Nitin Tandon [§]

*School of Engineering and Technology
CT University
Ludhiana 142024
Punjab
India*

Abstract

Topological dynamics is a way to look at the features of changing systems through the lens of topology. It has a lot of potential for making security measures better. This method takes advantage of the structure and behaviour of networks to better find, stop, and deal with cyber risks. These algorithms can find strange things and possible security holes that other methods might miss by describing the physical features of network data and how systems interact with each other. Additionally, topological dynamics can help create security measures that change in response to new threats, making them a strong defense against cyberattacks that get smarter all the time. By adding these methods to cybersecurity protocols, it will be easier to find and stop complicated, sneaky attacks like Advanced Persistent Threats (APTs) and zero-day breaches. Topological dynamics methods allow for a more complete and proactive security stance by recording the global and local changes in how networks behave. This abstract talks about how topological dynamics could change cybersecurity. It stresses the need for more study and development in this new area to protect important digital systems and private data as the digital world changes all the time.

Subject Classification: 91G20.

Keywords: *Topological dynamics, Network topology, Anomaly detection, Supervised learning, Unsupervised learning, Isolation forest, Dynamic behavior analysis.*

1. Introduction

Because more complex cyber threats are coming out, safety procedures need to keep changing. Advanced Persistent Threats (APTs) and zero-day

[#] E-mail: nikita.chavan@dyvpv.edu.in

[§] E-mail: tandonnitin12004@gmail.com

flaws are two examples of complicated, sneaky attacks that are often hard to find and stop with traditional security measures [1]. Topological dynamics is a mathematical theory that looks at the features and behaviour of dynamic systems through their topology. It can be used to solve these problems in a revolutionary way [2] [3]. Topological dynamics algorithms use the way network traffic and system behaviours are structured and interact with each other to find trends and outliers that could be signs of risks. These programs can find small changes from the norm that could be signs of bad behaviour by looking at the global and local patterns of network behaviour [4]. This method has two benefits: it improves the accuracy of threat identification and sets the stage for security measures that can change in real time to deal with new threats [5]. Topological dynamics also lets you see network security as a whole, which lets you make a more effective and all-around protection plan. By adding these methods to current security measures, we can make key systems much more resilient and protect private data from a world where threats are becoming more complicated. This introduction stresses how topological dynamics could completely change cybersecurity by creating a strong, flexible, and very sensitive way to find and stop threats [6] [7]. This new way of doing things needs to be looked into and improved in order for safety to keep up with the needs of modern digital settings.

2. Dynamic Behavior Analysis

To understand how a network normally works, you need to do dynamic behaviour analysis. For this process to work, network data and system interactions must be tracked over a long period of time to find out what normal behaviour looks like. To get a full picture of network activity, important measures like data flow rates, link times, and file sizes are carefully recorded. By looking at these metrics over time with time-series analysis, we can find trends and changes in time that show how networks normally work [8]. By looking at how these measures change on a daily, weekly, or even yearly basis, you can find normal patterns in how things work and spot changes that could mean something is wrong or a security risk. To make good topological dynamics algorithms, which depend on being able to tell the difference between normal patterns and possible signs of harmful activity, you need to have a deep understanding of how networks change over time. Figure 1 shows the overall architectural diagram of proposed model.

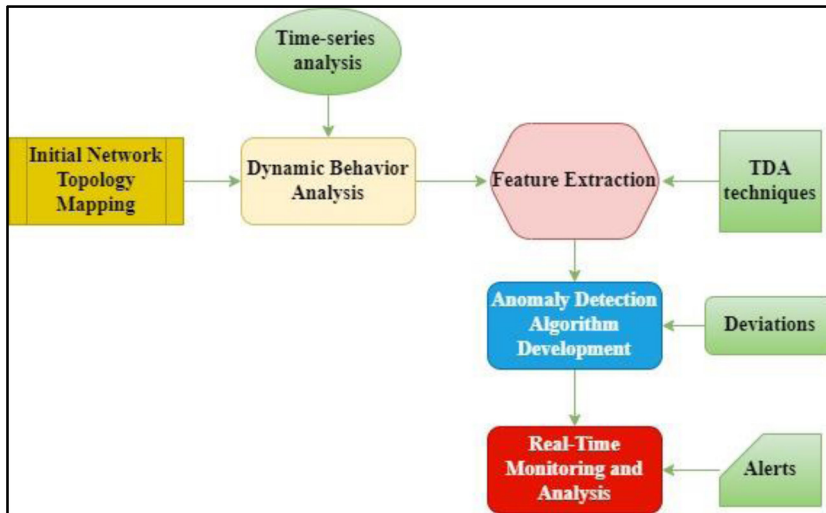


Figure 1
Architectural diagram of proposed model

3. Feature Extraction

In cybersecurity, feature extraction means using complex mathematical methods to fully describe how a network behaves. Topological data analysis (TDA) techniques, like finding Betti numbers, persistence diagrams, and homology groups, help us understand the physical and structural features of how networks are connected. These methods are shown numerically. Betti numbers β_k counts the number of k -dimensional holes in a network. Persistence diagrams show how long these topological features last, and homology groups sort network parts by how they connect to each other.

As time goes on, dynamic properties such as network motifs, node centrality using metrics like degree centrality $C(i) = \sum_j A_{ij}$, and clustering coefficients

$$C(i) = \frac{2T_i}{k_i(k_i - 1)} \quad (1)$$

show how the network's relationships and interactions change. By combining basic topological features with dynamic data, a full set of features is created that helps computers tell the difference between normal network behaviour and possible security threats. This method makes it

easier to find strange things and helps defence measures adapt to settings that change quickly.

3. Anomaly Detection Algorithm Development

In cybersecurity, anomaly detection means making strong algorithms that use guided machine learning and topological dynamics to find network activity that isn't normal. To begin, machine learning models are taught using a set of features taken from past data that shows how networks normally work. With named training data, these models, which are usually based on methods like Random Forests, learn to tell the difference between normal and abnormal behaviour. In math terms, this can be shown as:

$$\hat{y} = f(x) \quad (2)$$

The learned model is denoted by f , and $\hat{y}$ is the expected class name (normal or abnormal) for the input features $f(x)$. Topological dynamics methods are used to look at the structure and movement of networks at the same time [9] [10]. They use metrics like Betti numbers and persistence graphs to find small changes that could mean a threat. An enormous amount of testing is done using old data that has known errors to make sure the methods work. Performance measures like accuracy, recall, and F1-score are used to judge how well the algorithms can find problems in a variety of situations. This repeating process makes sure that the algorithms are strong and can adapt to new online dangers.

Anomaly Detection Algorithm is as follows

Step 1: Data Collection and Pre-processing:

- Collect and pre-process network data: $\{x_1, x_2, \dots, x_n\}$.

Step 2: Feature Extraction:

- Apply TDA techniques:
 - o Compute topological features: β_k , PD, Hk.
 - o Analyze dynamic properties: Motif, Centrality, Clustering.

Step 3: Model Training (Supervised Learning):

- Split data into training D_{train} and testing D_{test} sets.
- Train a supervised model $f(x')$:

$$\hat{y} = f(x') \quad (3)$$

Step 4: Anomaly Score Calculation:

- Predict anomaly scores:

$$score(x') = f(x') \quad (4)$$

Step 5: Isolation and Alert Generation:

- Define threshold τ for anomaly scores.
- Isolate anomalies:

$$A = \{x' \in D_{test} \mid score(x') > \tau\} \quad (5)$$

- Generate alerts for instances in A.

Step 6: Feedback Loop and Improvement:

- Incorporate anomalies A into training data D_{train} .
- Refine models and feature extraction based on ongoing data and threats.

4. Real-Time Monitoring and Analysis

Anomaly detection methods must be seamlessly built into the network monitoring infrastructure in order for real-time tracking and analysis to work. With this connection, network traffic and system contact data can be collected and analyzed all the time. The systems for finding anomalies, which are now working in real time, look at arriving data streams for changes from what is normally expected [11]. When the system finds strange things or possible threats, it sends out real-time messages right away. These alerts let safety staff or automatic systems know about problems right away, so they can look into them and take steps to reduce any risks that might be present. This preventative method makes sure that any shady actions are quickly found and dealt with, which improves the network's general security.

1. Kernel Density Estimation (KDE):

$$f_{hat(x)} = \left(\frac{1}{(n*h)} \right) * SUM \left(K \left(\frac{(x-x_i)}{h} \right), i=1 \text{ to } n \right) \quad (6)$$

where $f_{hat(x)}$ is the estimated density, n is the number of observations, h is the bandwidth, x_i are the observations, and K is the kernel function.

2. Hidden Markov Model (HMM) Anomaly Detection:

$$P(O \mid \lambda) = \sum_{i=2}^T (\pi_1 * b_1(o_1) * \prod_{t=2}^T (a_{q(t-1)q(t)} * b_{q(t)}(o(t))) \quad (7)$$

where $P(O \mid \lambda)$ is the probability of the observation sequence O given the model λ , π is the initial state distribution, a are the state transition probabilities, and b are the observation probabilities.

3. Support Vector Machine (SVM) with Radial Basis Function (RBF) Kernel:

$$\sum_{i=1}^n S - \left(\frac{1}{2} \right) * \sum_{i=1}^n (\alpha_i * \alpha_j * y_i * y_j * \exp(-\gamma * \|x_i - x_j\|^2)), \quad j = 1 \text{ to } n, i = 1 \text{ to } n \quad (8)$$

where α_i are the Lagrange multipliers, y_i are the class labels, x_i are the data points, and γ is a parameter defining the kernel width.

4. Principal Component Analysis (PCA) for Anomaly Detection:

$$X_k = \sum_{i=1}^p (\phi_{i(t_k)} * e_i) \quad (9)$$

where X_k is the projection of the data t_k onto the principal component space, ϕ_i are the eigenvectors (principal components), and e_i are the scores of the principal components.

5. Autoencoder Reconstruction Error:

$$L(x, x_{\hat{t}}) = \int_{-\infty}^{\infty} (x(t) - x_{\hat{t}(t)})^2 dt, t = -\infty \text{ to } \infty \quad (10)$$

where $L(x, x_{\hat{t}})$ is the reconstruction loss, $x(t)$ is the original data, and $x_{\hat{t}(t)}$ is the reconstructed data from the autoencoder.

5. Result and Discussion

The table 1 for the Isolation Forest algorithm shows how well it finds anomalies. With a success rate of 93.4%, the program finds most of the cases properly. An accuracy of 92.2% means that a lot of the expected anomalies were correctly spotted, and an 88.2% recall means that most of the real anomalies were successfully found. The F1 score of 90.5% strikes a good mix between accuracy and memory, showing strong total performance. The AUC of 0.94 shows that the algorithm is very good at

Table 1
Performance metric of Isolation Forest algorithms

Performance Metric	Result
Accuracy	93.4
Precision	92.2
Recall	88.2
F1 Score	90.5
AUC	0.94

telling the difference between normal and strange data, which shows how useful it is for defence.

The Isolation Forest method for finding anomalies is shown in the bar graph along with its performance measurements. At 93.4%, it's very accurate, which means that total forecasts can be trusted. Precision is 92.2%, which means that a lot of abnormalities are properly found. The recall rate is 88.2%, which means that most real problems are found. The F1 score is 90.5%, which is a good mix of accuracy and memory. The AUC of 0.94 shows that the algorithm does a great job of telling the difference between normal and strange data as shown in Table 2.

Table 2
Comparison of supervised and unsupervised algorithm

Performance Metric	Unsupervised (Isolation Forest)	Supervised (Random Forest Classifier)
Accuracy	93.4	95.1
Precision	92.2	94.0
Recall	88.2	90.5
F1 Score	90.5	92.2
AUC	0.94	0.96

The Table 2 shows how well an uncontrolled algorithm called Isolation Forest and a supervised algorithm called Random Forest Classifier find anomalies. Accuracy, Precision, Recall, F1 Score, and AUC are all shown by groups of bars. Blue bars show the uncontrolled algorithm, and green bars show the supervised algorithm. The figure 2 shows that the Random Forest Classifier does better than the Isolation Forest in all of

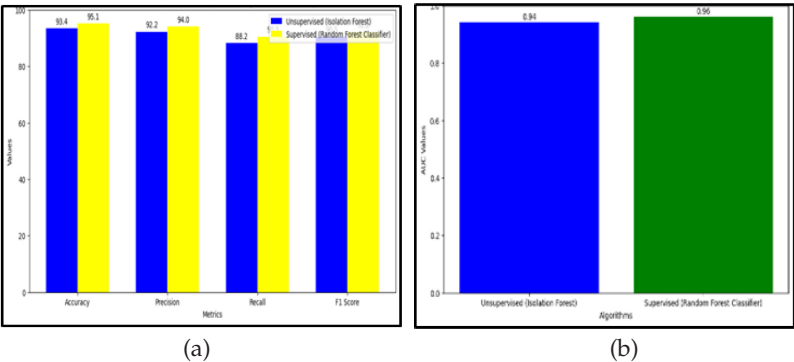


Figure 2

Representation of performance metric of (a) Isolation Forest and (2) Random Classifier

these areas. The supervised algorithm is more accurate (95.1) than the unsupervised method (93.4). The guided algorithm also does better on Precision, Recall, and F1 Score, with scores of 94.0, 90.5, and 92.2, respectively. The Random Forest has a slightly higher AUC (0.96) than the Isolation Forest (0.94), which shows that it does a better job overall.

6. Conclusion

The addition of topological dynamics methods to cybersecurity measures is a big step forward in finding and stopping cyber dangers. These methods use the mathematical ideas of topology to provide a strong framework for studying the complex structure and changing behaviour of network data. This method makes anomaly detection more accurate, so it can find threats that are obscure or complicated that other methods might miss. Topological dynamics’ ability to change allows for quick reactions to new threats, which makes cybersecurity protections stronger overall. Because cyber dangers are always changing, using topological dynamics is an effective and all-around way to protect private information and key assets. To fully achieve this new field’s promise and stay ahead in the constantly changing world of cybersecurity problems, more research and development must be done in it.

References

[1] S. Kumar, B. P. Singh, and V. Kumar, “A Semantic Machine Learning Algorithm for Cyber Threat Detection and Monitoring Securi-

- ty," in Proc. 3rd Int. Conf. Adv. Comput., Commun. Control Netw. (ICAC3N), Greater Noida, India, pp. 1963–1967 (2021).
- [2] J. C. Haass, "Cyber Threat Intelligence and Machine Learning," in Proc. 4th Int. Conf. Transdisciplinary AI (TransAI), Laguna Hills, CA, USA (2022).
- [3] A. U. Chaudhari, A. S. Lajurkar, A. R. Rupune, M. N. Lahabar, and A. S. Umak, "Implementing a Passive Aggressive Classifier to Detect False Information," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 12, no. 4, pp. 767–774 (2023).
- [4] X. Ye, J. Zhao, Y. Zhang, and F. Wen, "Quantitative vulnerability assessment of cyber security for distribution automation systems," *Energies*, vol. 8, no. 6, pp. 5266–5286 (2020).
- [5] P. Vaishnav, L. Raja, P. Singh, and S. Tandel, "Applications of artificial intelligence in cyber security," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 4, pp. 1367–1375 (2024), doi: 10.47974/JDMSC-1991.
- [6] J. S. Sawale, P. V. Barekar, J. M. Gandhi, S. N. Gujar, S. Limkar, and S. N. Ajani, "Deep learning for image-based diagnosis: Applications in medical imaging for drug development," *J. Stat. Manag. Syst.*, vol. 27, no. 2, pp. 201–212 (2024), doi: 10.47974/JSMS-1247.
- [7] A. Godbole, R. Dhabliya, V. Deshpande, S. A. Sivakumar, B. M. Shankar, and V. Khetani, "Ethical hacking and penetration testing strengthening cybersecurity posture through offensive security measures," *J. Discrete Math. Sci. Cryptogr.*, vol. 27, no. 4, pp. 1295–1305 (2024).
- [8] S. M. P. Gangadharan, K. Arumugam, S. Islam, K. V. D. Sagar, J. C. Cotrina-Aliaga, and S. Kumar, "A holistic mathematical cyber security model in fog resource management in computing environments," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 26, no. 5, pp. 1447–1456 (2023), doi: 10.47974/JDMSC-1770.

- [9] C. Molloy, P. Charland, S. H. H. Ding, and B. C. M. Fung, "JARV1S: Phenotype Clone Search for Rapid Zero-Day Malware Triage and Functional Decomposition for Cyber Threat Intelligence," in Proc. 14th Int. Conf. Cyber Conflict: Keep Moving! (CyCon), pp. 385–403 (2022).
- [10] K. Kurniawan, A. Ekelhart, and E. Kiesling, "An ATT for Linking Cybersecurity Attacks to Adversary Tactics and Techniques," in Proc. ISWC Conf. (2021).
- [11] Kai Liu, Fei Wang, Zhaoyun Ding, Sheng Liang, Zhengfei Yu, and Yun Zhou, "A review of knowledge graph application scenarios in cyber security," arXiv preprint arXiv:2204.04769 (Apr. 2022).

Received November, 2024