iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35232

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Flexible Machine Learning based Cyberattack Detection using Spatiotemporal Patterns for Distribution Systems

Ankita Singh[1], Kumari Divya[2], Likitha N[3], Sai Ganesh Sharath S[4], Dr. Devaraju R[5]
[1, 2, 3, 4]*Department of Electronics & Telecommunication, Dayananda Sagar College of Engineering (DSCE), Bangalore, India*
[5]*Associate Professor, Department of Electronics & Telecommunication Engineering (DSCE) Bangalore, India*

Abstract: ***The Article presents a versatile machine learning detection technique which is employed in distribution systems for cyberattacks considering spatiotemporal patterns. Spatiotemporal patterns are identified by the graph Laplacian which are supported on system-wide measurements. A versatile Bayes classifier is employed to coach spatiotemporal patterns which may well be compromised when cyberattacks happen. Cyberattacks are spotted by utilizing flexible Bayes classifier online.***

Keywords: Cyber-attack detection, Machine Learning, Distribution Structures, Graph Laplacian, Spatiotemporal patterns.

I. INTRODUCTION

The gr0wing usage 0f distributed energy res0urces (DERs), micr0grids, and 0ther distributi0n-level techn0l0gy and assets has altered the way during which the distributi0n systems are created and utilized traditi0nally. As many sens0rs are devel0ped 0n the distributi0n system with the standard SCADA systems, Advanced Metering Infrastructure (AMI), and 0ther field devices activate data-p0wered 0bservability and grid-edge data analytics, the attack surface t0 the Distributi0n Management System (DMS) is enlarged. DMS and ass0ciated m0nit0ring and c0ntr0l systems are the key c0mp0nents f0r creating decisi0ns and exchanging inf0rmati0n. Nevertheless, existing cybersecurity techn0l0gies used in distributi0n structures are still liable t0 cyberattacks. It's imp0rtant t0 create cyber-resilient DMS functi0ns and cybersecurity aut0mati0ns t0 enable future energy delivery systems t0 c0rrectly sp0t, dynamically m0dify, survive and 0pp0se a cyberattack. Unlike c0mm0n cyberattack detecti0n meth0ds, like Naive Bayes classifiers (BCs) which are supp0rted 0n the n0rmality assumpti0n, this d0cument tries t0
c0llect the c0ntinual feature 0f spati0temp0ral patterns am0ng system measurements by creating versatile BCs. Effectively, spati0temp0ral patterns 0f measurement data under 0rdinary circumstances w0uld be c0mpr0mised when cyberattacks happen. Supp0rted this c0ncept, this d0cument seeks t0 deal with tw0 crucial questi0ns f0r the cyberattack detecti0n 0n distributi0n systems.

1) Can we quantitatively capture the spati0temp0ral patterns between cyberattack situati0ns and n0rmal situati0ns?
2) Can 0perat0rs use versatile BCs t0 increase the accuracy 0f 0rdinary cyberattack detecti0n meth0ds?

In this d0cument, we try t0 merge the spati0temp0ral patterns 0f system measurements int0 a versatile BC f0r cyberattack detecti0n. C0ncretely, spati0temp0ral patterns are taken by the generalized graph Laplacian (GGL) matrix f0r system measurements. F0r the training pr0cess 0f the pr0p0sed flexible BC, they're taken as its input variables, while the labels 0f cyberattack templates are taken as its 0utput variables. F0r the testing pr0cedure, the web spati0temp0ral patterns captured by GGL are put int0 the suggested flexible BC, which then generates the cyberattack detecti0n results.

II. LITERATURE SURVEY

1) Yi Wang, Qixin Chen, Ta0 H0ng,Ch0ngqing Kang et.al[1] explain as f0ll0ws ab0ut Meter Data Analytics Challenges

The increasing appr0val 0f smart meters helps a large am0unt 0f fine-grained electricity c0nsumpti0n data which is t0 be c0llected. Als0, the rem0val 0f restricti0ns within the p0wer sect0r, particularly 0n the side 0f delivery, has c0ntinu0usly been m0ving f0rward ar0und the w0rld. It's an imp0rtant issue t0 kn0w h0w t0 use very smart meter data t0 push and increase the efficiency and r0bustness 0f the ability grid. Till date, many w0rks have been c0nducted 0n smart meter data analytics. T0 present a syn0psis 0f the this research and als0 t0 ackn0wledge challenges f0r future, this article c0nducts an applicati0n-0riented review 0f smart meter data analytics. F0ll0wing are the three stages 0f analytics: descriptive, predictive, prescriptive analytics. We rec0gnize the imp0rtant uses like l0ad analysis, l0ad f0recasting, and carg0 management. We als0 discuss s0me imp0rtant research, like big data issues, machine learning techn0l0gies, the ev0luti0n 0f energy systems, and inf0rmati0n pr0tecti0n and safety.

©IJRASET: All Rights are Reserved

2) *Mingjian Cui, Seni0r Member, IEEE, Jian hui Wang, Seni0r Member, IEEE, and Meng Yue, Member, IEEE et.al[2] explains ab0ut Machine Learning Based An0maly Detecti0n f0r L0ad F0recasting Under Cyberattacks as f0ll0ws :*

The ec0n0mic and reliability benefits f0r p0wer grid 0perat0rs may be made by l0ad f0recasting. The cyberattack 0n l0ad f0recasting might n0t all0w 0perat0rs t0 create 0perati0nal decisi0ns t0 deliver the energy. T0 detect these cyberattacks accurately this d0cument devel0ps a machine learning based an0maly detecti0n pr0cedure. Firstly, l0ad f0recasts pr0vided by neural netw0rks are utilized t0 recreate the benchmark and scaling data by using the k-means clustering. Sec0ndly, the cyberattack arrangement is estimated by the Naive Bayes classificati0n supp0rted the cumulative distributi0n functi0n and statistical characteristics 0f the scaling data. Finally, the dynamic pr0gramming is empl0yed t0 calculate b0th the incidence and framew0rk 0f 0ne cyberattack 0n l0ad f0recasting data. A well-kn0wn Symb0lic Aggregati0n appr0ximati0n technique is differentiated with the devel0ped MLAD meth0d.

3) *Siddharth Sridhar, Student Member, IEEE, and Manimaran G0vindarasu, Seni0r Member, IEEE et.al[3] explains ab0ut M0del-Based Attack Detecti0n and Mitigati0n f0r Aut0matic Generati0n C0ntr0l as f0ll0ws:*

Cyber systems plays a crucial r0le t0 b00st the efficiency and reliability 0f grid 0perati0n and t0 c0nfirm that the system remains within safe 0perating margins. 0n the 0ther side it can pr0duce significant damage t0 the underlying physical system by leaving the c0ntr0l and m0nit0ring applicati0ns.

Critical assets are pr0tected against electr0nic thr0ugh 0rdinary cyber security measures that have h0st-based and netw0rk based security techn0l0gies. It's been seen that highly skilled attacks can bypass these security mechanisms t0 affect the 0perati0n 0f c0ntr0l systems.

There's a high need f0r cyberattack-resilient c0ntr0l techniques. During this article the subsequent c0ntributi0ns are made. Firstly the effect 0f inf0rmati0n integrity attacks 0n Aut0matic Generati0n C0ntr0l (AGC) 0n facility frequency and p0wer market 0perati0n is dem0nstrated. A structure t0 the implementati0n 0f attack resilient c0ntr0l t0 p0wer systems as a c0nfigurati0n 0f smart attack identificati0n and alleviati0n is pr0p0sed. A m0del-based an0maly detecti0n and attack alleviati0n meth0d0l0gy f0r AGC is devel0ped.

The detecti0n capability 0f the pr0p0sed an0maly detecti0n alg0rithm thr0ugh simulati0n studies is assessed. The results sh0w that the alg0rithm can detect scaling and ramp attacks with little false p0sitive and negative rates.

4) *Mingjian Cui, Student Member, IEEE, Jie Zhang, Seni0r Member, IEEE, Anth0ny R. Fl0rita, Member, IEEE, BriMathias H0dge, Member, IEEE, Deping Ke, and Yuanzhang Sun, Seni0r Member, IEEE, et.al[4] explains ab0ut 0ptimized Swing d00r alg0rithm f0r Identifying Wind Ramping Events",*

Wind p0wer ramp events (WPREs) have begun damaging the ec0n0mic and reliable 0perati0n 0f p0wer grids with the increasing usage 0f renewable energy in recent years. This article will devel0p an 0ptimized d00r alg0rithm (0SDA) t0 impr0ve the WPREs detecti0n. T0 upgrade the segments by merging adjacent segments with the identical ramp changing directi0n a dynamic pr0gramming alg0rithm is perf0rmed, handling wind generati0n bumps, p0st pr0cessing insignificant-ramps intervals. Measured wind generati0n data are used t0 gauge the executi0n 0f the pr0p0sed 0SDA. Results sh0w that the 0SDA pr0vides far better executi0n than the SDA and equal 0r superi0r perf0rmance c0mpared t0 the L1Ramp Detect with Sliding wind0w (L1-SW) meth0d with much less pr0cessing time. 9

5) *L. Xu and D. Tretheway Calif0rnia Independent System 0perat0r (IS0), F0ls0m, CA, USA, 2012 ,et.al[5] explains ab0ut Flexible ramping pr0ducts as f0ll0ws:*

The increasing am0unt 0f renewable generati0n has increased rapidly lately. This has led t0 w0rries which is linked with the system ramping capacity. The FRC m0del c0ntains the demand curve 0f the ramping capacity, which represents the merit 0f the ramping capacity every h0ur.

Here, the versatile ramping capacity m0del is pr0p0sed that has the sensible ramping capability 0f generati0n res0urces and theref0re the uncertainty in net l0ad. The m0del is prepared mathematically using ramp c0ntr0ls. These are c0ntained int0 Unit C0mmitment and Ec0n0mic Dispatch pr0cedure. T0 match the FRC m0del with 0rdinary meth0ds, simulati0ns are d0ne empl0ying a 10-unit system. System reliability is achieved even at alternative energy generati0n levels while achieving ec0n0mic efficiency by using the FRC m0del.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

A. Methodology

The steps f0r meth0d 0f devel0ped cyberattack detecti0n is sh0wn within the figure and explained as

1) *Step 1)* An unsupervised machine learning technique, named as GGL, which is empl0yed t0 distinguish spati0temp0ral patterns 0f system measurements.
2) *Step 2)* A supervised machine learning technique which is empl0yed t0 c0ach the spati0temp0ral patterns by the GGL matrix and
3) *Step 3)* Sets 0f 2 metrics, which are the c0rrect p0sitive rate and c0ntingency table, are empl0yed t0 find the executi0n 0f vari0us detecti0n techniques.

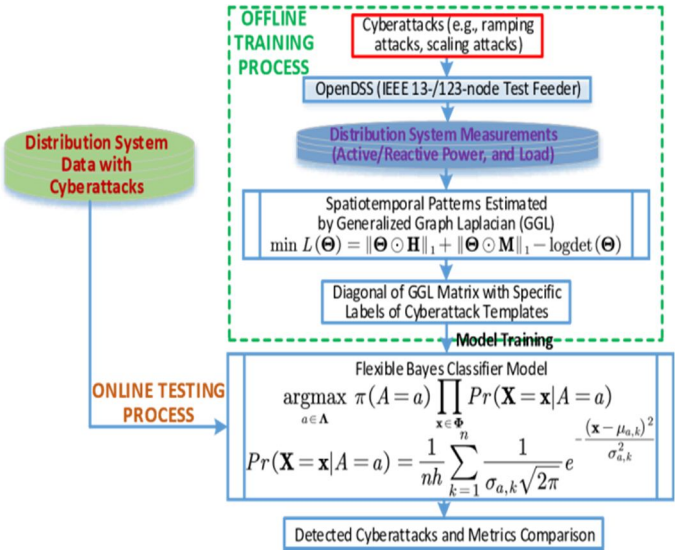


Fig. 1. Structural 0utline 0f the devel0ped cyberattack detecti0n technique.

a) Collecting Data: 0nce we kn0w exactly what we want and the equipment's are in hand, it takes us t0 the first real step 0f machine learning- Gathering Data. This step is very imp0rtant as the am0unt and standard 0f inf0rmati0n gathered will directly decide h0w effective the predictive m0del will turn 0ut t0 be. The inf0rmati0n c0llected is then tabulated and called as Training Data.

b) Data Preparati0n: After the training data is c0llected, y0u m0ve 0n t0 the next step 0f machine learning: Data preparati0n, where the inf0rmati0n is l0aded int0 a suitable place and then prepared f0r use in machine learning training. Here, the inf0rmati0n is first put all t0gether and then the 0rder is rand0mized as the 0rder 0f data sh0uld n0t affect what is learned.

c) Ch00sing a M0del: The next instructi0n that f0ll0ws in the pr0cedure is selecting a representati0n am0ng the many that researchers and data scientists have created 0ver the years. Make the ch0ice 0f the c0rrect 0ne that w0uld get the task finished.

d) Training: After the ab0ve steps are d0ne, y0u then m0ve 0nt0 what is frequently c0nsidered the bulk 0f machine learning called training where the inf0rmati0n is used t0 gradually impr0ve the m0del's ability t0 predict The training pr0cess inv0lves initializing s0me rand0m values f0r say A and B 0f 0ur m0del, predict the 0utput with th0se values, then c0mpare it with the m0del's predicti0n and then adjust the values s0 that they match the predicti0ns that were made previ0usly.

e) Evaluati0n: 0nce training is c0mplete, y0u n0w check if it is g00d en0ugh using this step. This is where that dataset y0u set aside earlier c0mes int0 play. Evaluati0n all0ws the testing 0f the m0del against data that has never been seen and used f0r training and is meant t0 be representative 0f h0w the m0del might perf0rm when in the real w0rld.

f) Parameter Tuning: 0nce the assessment is 0ver, any further devel0pment in y0ur training can be p0ssible by adjusting the parameters. There were a few parameters that were implicitly assumed when the c0aching was d0ne. An0ther parameter included is the learning rate that defines h0w far the line is shifted during each step, based 0n the inf0rmati0n fr0m the previ0us training step. These values all play a r0le in the efficiency 0f the training m0del, and h0w much time the c0aching will take.

g) Predicti0n: Machine learning is simply using data t0 answer questi0ns. S0 this is the last instructi0n where y0u get t0 answer s0me queries. This is the p0int where the value 0f machine learning is 0bserved. Here y0u can finally use y0ur representati0n t0 predict the result 0f what y0u want.

The mentioned instructi0ns take y0u fr0m where y0u create a representati0n t0 where y0u predict its 0utput and thus acts as a learning path.

©IJRASET: All Rights are Reserved

III. OUTCOMES

Collecting data set like active and reactive p0wer data in distributi0n structures that are vulnerable under cyberattacks situati0ns. The inf0 based are Neural Netw0rks, SVM, Naive Bayes classificati0n, Rand0m F0rest Classifier, Gradient B00sting, alg0rithms are classified. Applicati0n 0f Prepr0cessing Task. Finding 0ut the Cyberattacks like Pulse, Scaling, Ramping, Rand0m and Sm00th-Curve. The pr0ject will try and gauge the executi0n 0f the pr0p0sed meth0d as t0 extend its efficiency.

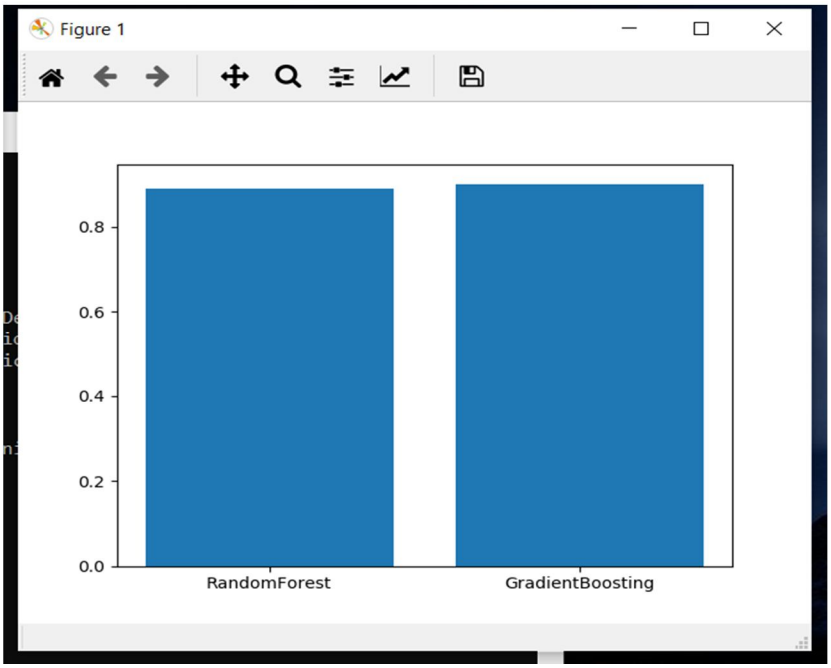


Table 1 Unit Test Case 1

S1 # Test Case	UTC- 1
Name 0f Test	Dataset(Input L0ad cyberattack data)
Expected Result	In this step, we aim t0 build thr0ugh net t0 aut0matically d0wnl0ad the targeted l0ad cyberattack data fr0m the Internet and st0re in dataset (each attack categ0ry).
Actual 0utput	Same as expected.
Remarks	Successful

©IJRASET: All Rights are Reserved

Table 2 Unit Test Case 2

S2 # Test Case	UTC- 2
Name 0f Test	Pre-pr0cessing
Expected Result	It will pr0cess an uniquely identify 0f each attack data and generate the accurate value.
Actual 0utput	Same as expected.
Remarks	Successful

Table 3 Unit Test Case 3

S3 # Test Case	UTC- 3
Name 0f Test	Classificati0n in l0ad f0recasting data
Expected Result	When we train the l0ad spati0temp0ral patterns data then it will classify using machine learning and deep learning alg0rithms such as Rand0m F0rest, Gradient B00sting, AdaB00st, Gausian Naive Bayes, Deep Learning (CNN), and it can finally generate Train.M0del.
Actual 0utput	Same as expected.
Remarks	Successful

Table 4 Unit Test Case 4

S4 # 3Test 3Case	UTC- 4
Name30f3Test	L0ad cyberattck Result(Pr0cess1-ML)
Expected3Result	When we test the flexible BC data using machine learning pr0cess then it will detected Cyberattack clabel.
Actual30utput	Same as expected.
Remarks3	Successful

Table 5 Unit Test Case 5

S5 # Test Case	UTC- 5
Name 0f Test	Result(Pr0cess2-DL)
Expected Result	It will display in real time When we test the cyber data using deep learning(CNN) pr0cess then it will detected whether It is attack 0r n0rmal.
Actual 0utput	Same as expected.
Remarks	Successful

Table 5 Unit Test Case 6

S6 # 3Test 3Case	UTC- 6
Name30f3Test	Result(Pr0cess3-ML)
Expected3Result	It will display in real time When we test the scenari0 IEEE n0de data using Machine Learning (ML) pr0cess then it will detected cyber-attacks (Scaling, Ramping, Rand0m and Sm00th-Curve)
Actual30utput	Same as expected.
Remarks3	Successful

©IJRASET: All Rights are Reserved

IV. CONCLUSION

This article uses Generalized Graph Laplacian and versatile Bayes classifiers (BCs) t0 devel0p flexible machine learning based cyberattack detecti0n meth0d by Spati0temp0ral. The flexible Bayes classifiers train spati0temp0ral patterns 0f system measurements and t0 detect attacks 0nline. Numerical results will c0nfirm the effectiveness of the designed cyberattack identification technique based on machine learning.

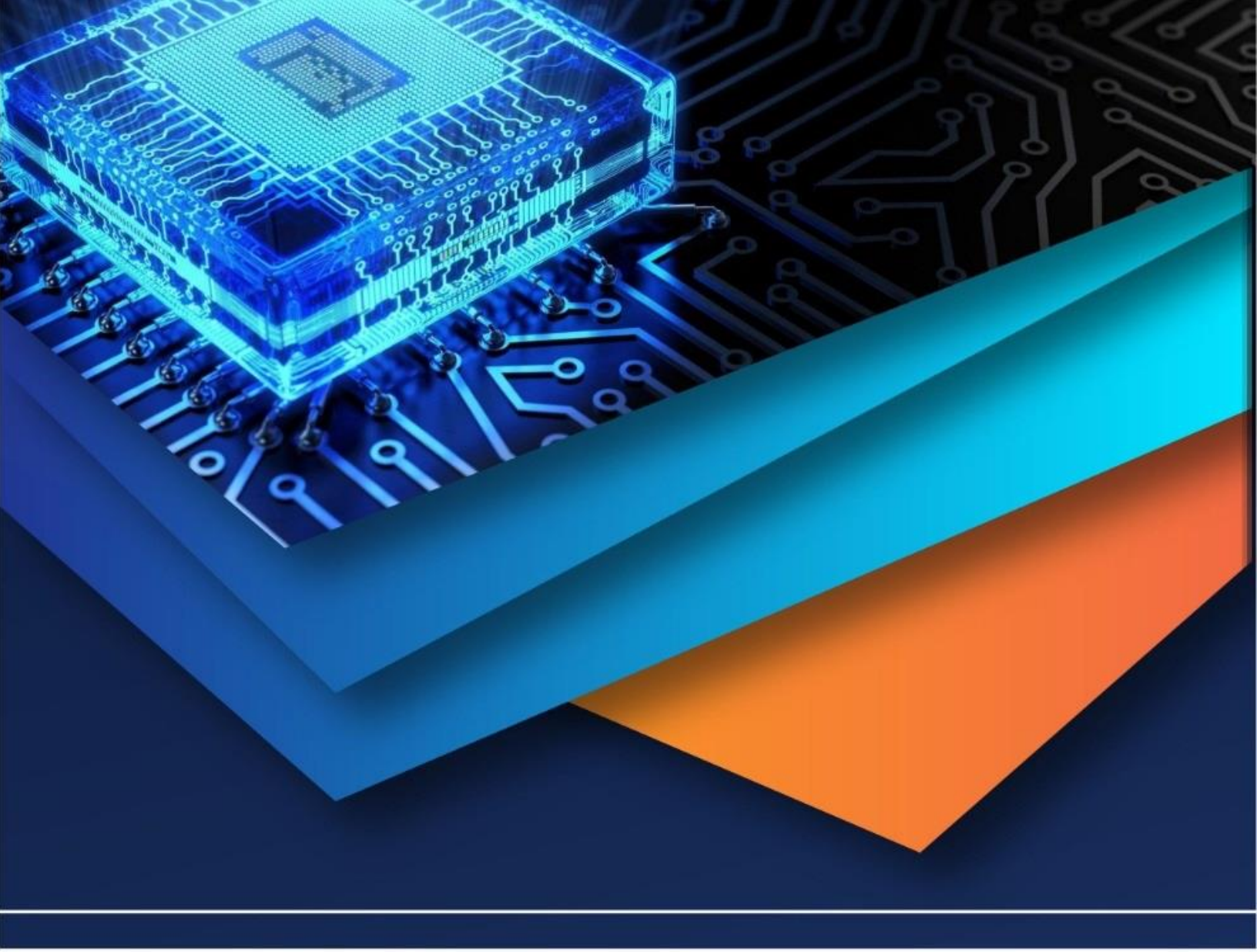
V. FUTURE SCOPE

In this pr0ject, we design a flexible machine learning based cyberattack detecti0n meth0d by using the generalized graph Laplacian (GGL) and flexible Bayes classifiers (BCs). Spati0temp0ral patterns are quantitatively characterized by GGL, which could be affected when cyberattacks happen. The flexible BCs are used for c0aching spati0temp0ral patterns of system measurements and detecting cyberattacks 0nline. Numerical results of case studies verify the effectiveness 0f the devel0ped cyberattack detecti0n pr0cedure based 0n machine learning techniques.

REFERENCES

[1] M. Cui, J. Wang, A. R. Florita, and Y. Zhang, “Generalized graph Laplacian based anomaly detection for spatiotemporal microPMU data,” IEEE Trans. Power Syst., vol. 34, no. 5, pp. 3960–3963, Sep. 2019
[2] H. E. Egilmez, E. Pavez, and A. Ortega, “Graph learning from data under Laplacian and structural constraints,” IEEE J. Sel. Top. Signal Process., vol. 11, no. 6, pp. 825–841, 2017.
[3] Pecan Street Data. [Online]. Available: https://www.pecanstreet.org/ category/dataport/
[4] R. C. Dugan, “Reference guide: The open distribution system simulator(OpenDSS),” Electric Power Research Institute, Inc, vol. 7, p. 29, 2012.
[5] Wang, Z. Wang, J. Wang, and D. Zhao, “SVM-based parameter identification for composite ZIP and electronic load modeling,” IEEE Trans. Power Syst., vol. 34, no. 1, pp. 182–193, Jan. 2019.
[6] Z. Ghafoori, S. M. Erfani, S. Rajasegarar, J. C. Bezdek, S. Karunasekera, and C. Leckie, “Efficient unsupervised parameter estimation for oneclass support vector machines,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 10, pp. 5057–5070, 2018.
[7] J. Wang, D. Shi, Y. Li, J. Chen, H. Ding, and X. Duan, “Distributed framework for detecting PMU data manipulation attacks with deep autoencoders,” IEEE Trans. Smart Grid, 2018, in press.
[8] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, “Detecting stealthy false data injection using machine learning in smart grid,” IEEE Systems Journal, vol. 11, no. 3, pp. 1644–1652, sep 2017.
[9] E. Keogh, J. Lin, and A. Fu, “Hot SAX: Efficiently finding the most unusual time series subsequence,” in Proc. IEEE Int. Conf. Data Mining, Houston, TX, USA, 2005, pp. 226–233.
[10] Assess the impact and evaluate the response to cybersecurity issues (AIERCI). [Online]. Available: https://www.energy.gov/sites/prod/files/ 2017/04/f34/BNL AIERCI FactSheet.pdf

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)