Robust Network Anomaly Detection with K-Nearest Neighbors (KNN) Enhanced Digital Twins

Peprah Obed Adjei, Sumit Kumar Tetarave, Caroline John, Madlyn Manneh, and Parthasarathi Pattnayak

Abstract—**Modern network security remains a critical concern in the digital landscape due to evolving cyber threats and increasingly sophisticated attack vectors such as Advanced Persistent Threats and Zero-Day Vulnerabilities. Leveraging advanced technologies such as artificial intelligence (AI) and machine learning (ML) can enhance threat detection capabilities and improve incident response times when detecting and mitigating network security threats. On the other hand, an imbalanced dataset of network traffic in AI/ML models presents several challenges and can significantly impact the performance and effectiveness of the models to predict attacks. Our research aims to amplify the robustness of the imbalanced network traffic dataset to fit the analysis and adaptability of KNN-based Digital Twins dedicated to network anomaly detection. This paper capitalizes on the remarkable performance of the model, characterized by impeccable precision, recall, and F1-score, as indicated by the classification report with 99% accuracy. The confusion matrix further highlights the model's performance using the proposed robustness dataset, showing a minimal False Positive Rate (FPR) compared to similar works in the literature.**

Index Terms—**Network Anomaly Detection; K-Nearest neighbors; Digital Twins; Robustness; Adaptability.**

I. Introduction

In an age where the digital realm intertwines seamlessly with the physical, ”Digital Twins” emerges as a beacon of technological innovation. A digital twin is a virtual representation of a real-world system or thing, breathing life into data and ushering in a new era of understanding, monitoring, and optimizing real-world entities [1]. Digital Twins' applications span various domains, offering transformative insights and capabilities. Digital Twins act as a dynamic lens in network management, enabling us to peer into the intricate web of interconnected devices, data flows, and protocols. They are not just replicas but sentient observers, change catalysts, and network integrity guardians [2].

Let us contemplate a situation where an Internet Service Provider (ISP) implements Digital Twins to mirror its expansive network infrastructure. These Digital Twins serve as vigilant sentinels, mirroring network nodes, devices, and the ebb and flow of data. By meticulously modeling and analyzing their digital counterparts, ISPs gain the foresight to preemptively identify congestion points, optimize routing, and enhance Quality of Service (QoS). The applications extend even further, delving into the realm of cybersecurity. Digital Twins become watchful protectors, embodying the network's normal behavior. They are the front line defenders, swiftly detecting and mitigating deviations, such as Distributed Denial of Service (DDoS) attacks.

Network anomalies are enigmatic deviations from the expected norm of network behavior. These deviations manifest in various forms, including unexpected surges in data traffic, aberrant bandwidth consumption, unauthorized access attempts, or erratic patterns in the transmission of network packets. For example, a sudden increase in data transfer rates during non-peak hours could indicate the beginning of a potential Distributed Denial of Service (DDoS) attack. In contrast, a consistent decrease in data transmission rates might suggest an underlying network issue [3].

These diverse and intricate anomalies pose a formidable challenge to network administrators and cybersecurity professionals. The ability to swiftly detect, analyze, and mitigate anomalies is paramount for safeguarding network integrity and mitigating potential threats. Thus, the overarching objective is to imbue Digital Twins with the intelligence to identify and respond effectively to these anomalies, thereby ensuring the steadfastness and security of network infrastructures [4]. Intriguingly, the fusion of Digital Twins with the power of machine learning, particularly K-nearest neighbors (KNN), ushers in a new frontier of network management [5].

This research explores the complexities of network anomalies, clarifying how they appear and the resulting outcomes. We explore this symbiotic relationship with a specific focus on harnessing KNN for network anomaly detection. In doing so, we aim to fortify the resilience and adaptability of Digital Twins in identifying and responding to network anomalies with precision on balanced and imbalanced network traffic datasets. The database used for this research was obtained from UNSW-NB15 Dataset ([6], [7]). It consists of 49 attributes, including network source IP (*srcip*), source byte (*sbyte*), attack categories (*attack_cat*), service, stat, and 135 protocols (*proto*) such udp, tcp, arp, cbt, nvp, ipv6-opts, etc. The evaluation shows that our balanced dataset mechanism's performance in identifying anomalies was enhanced significantly compared to existing ones [18].

The rest of the paper is organized as follows. Section II focuses on notable research in the relevant field. Section III explains the methods for acquiring the required findings alongside the model and proposed technique. Section IV presents the results and examines the proposed methodology. The conclusion of the research is presented in Section V, which also outlines the direction for future research.

Peprah Obed Adjei, Sumit Kumar Tetarave, Madlyn Manneh, and Parthasarathi Pattnayak are with the School of Computer Applications, Kalinga Institute of Industrial Technology, Bhubaneswar 751024, India e-mail: (nanakwameadjeipeprah@gmail.com, sumitkumar.fca@kiit.ac.in, Madlynmanneh03@gmail.com, and parthakiit19@gmail.com).

Caroline John is with Department of Cybersecurity, University of West Florida, Florida 32514, USA e-mail: cjohn@uwf.edu.

SoutheastCon 2024 | 979-8-3503-1710-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/SOUTHEASTCON52093.2024.10500053

979-8-3503-1710-7/24/$31.00 ©2024 IEEE

II. Literature Review

In recent years, anomaly detection has become a vital area of research in various applications. Several machine learning algorithms have gained prominence in this context [8] and highlighted the prevalence of specific algorithms frequently employed for anomaly detection. These techniques include Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forests (RF).

However, it is essential to acknowledge that adopting complex models rooted in Deep Learning (DL) architectures and ANN comes with inherent challenges. These challenges primarily revolve around the necessity for extensive training datasets, as emphasized by [9]. Additionally, models built upon such frameworks tend to exhibit low physical interoperability of their parameters, which can hinder their practical utility in specific applications. Consequently, alternative algorithms have been explored to address these challenges and enhance the effectiveness of anomaly detection. Of these choices, Random Forests (RF), One-Class Support Vector Machines (OCSVM), and Kernel Principal Component Analysis (KPCA) have established themselves as reputable and successful techniques [10]. Authors in [11] introduced an ML-based Digital Twin (DT) as a pivotal tool for efficiently managing real-world networks. This innovative approach enables network operators to design optimization solutions, troubleshoot network issues, perform simulations, and plan improvements effectively. The key to its effectiveness lies in utilizing deep learning techniques for modeling the DT. These techniques consider various network parameters, including traffic patterns, network topology, routing strategies, and scheduling policies. Consequently, the DT generates valuable performance metrics such as network utilization, latency, and packet loss. The iterative feedback process established between the DT and network optimization tools further enhances network configuration to meet network operator requirements, ensuring optimal network performance.

Despite their many advantages, these features also make DTs potential targets for security breaches. Unauthorized access to DTs can lead to misuse, a scenario referred to as the abuse case of DT [12]. Attackers can exploit their deep understanding of physical processes and devices accessible through DTs. This exploitation typically involves a two-stage strategy: first, manipulating DTs into a malicious state—altering the key data acquisition and dissemination processes, and then using this compromised state to manipulate the underlying physical system's behavior covertly. The reverse strategy, targeting Cyber-Physical Systems (CPS) to attack DTs, is also possible. For instance, the evolution of specialized malware tailored for Industrial Control Systems (ICS) indicates that adversaries possess in-depth knowledge of physical industrial processes [13]. With this understanding, attackers can infer and construct their knowledge about DTs, potentially leading to cascading failures.

The authors of ([14], [15]) have introduced the CPS digital-twin framework to overcome the cascading failures. This framework automatically empowers users to generate digital twins from CPS specifications, often represented as Automation ML-based engineering artifacts. These digital twins operate within a virtual environment. They can function independently of their physical counterparts, such as testing or closely emulating their program states to virtually replicate actual CPS behavior on the logic and network layers, primarily for monitoring purposes [16]. Moreover, the authors have emphasized that the digital twin concept offers numerous security-enhancing possibilities.

In the paper [17], the authors present a digital twin modal using KNN and MSVM machine learning algorithms. KNN achieved an accuracy of 86%. Moreover, authors in [18] implemented an ensemble intrusion detection technique for protecting the network traffic with 98.97% accuracy over HTTPS data sources. This method has a low false positive rate (FPR) of 2.58%. Our focus in this paper is to develop Intrusion Detection Systems (IDSs) to harness KNN features to improve accuracy and provide a mechanism to build a robust, balanced dataset that increases the effectiveness of identifying malicious links.

III. Proposed Methodology

We aim to construct a Digital Twin system capable of replicating network instances and promptly identifying deviations or anomalies by leveraging KNN in conjunction with our proposed data balancing method. This proactive approach will be pivotal in ensuring the continual enhancement of network security, aligning with the primary goal of our proposed work. Fig. 1 shows the proposed flow to accomplish the proposed objectives.

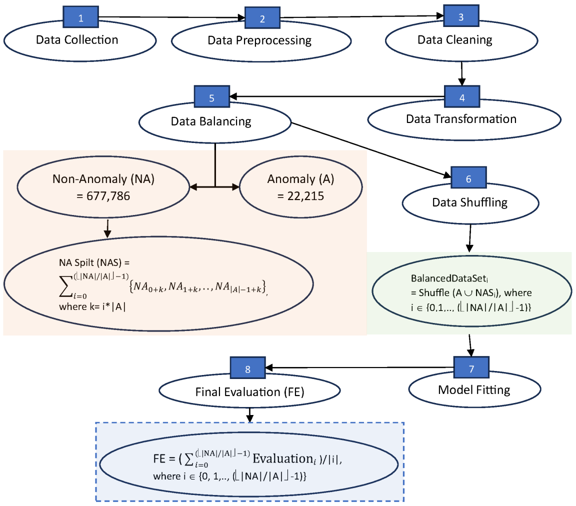

1
Data Collection
2
Data Preprocessing
3
Data Cleaning
4
Data Transformation
5
Data Balancing
Non-Anomaly (NA) = 677,786
Anomaly (A) = 22,215
NA Spilt (NAS) = $\sum_{i=0}^{(\lfloor |NA|/|A| \rfloor -1)} \{NA_{0+k}, NA_{1+k}, .., NA_{|A|-1+k}\}$,
where k= i*|A|
6
Data Shuffling
$BalancedDataSet_i$ = Shuffle (A ∪ NAS_i), where i ∈ {0,1,..., (⌊|NA|/|A|⌋ -1)}
7
Model Fitting
8
Final Evaluation (FE)
FE = ($\sum_{i=0}^{(\lfloor |NA|/|A| \rfloor -1)} Evaluation_i$)/|i|,
where i ∈ {0, 1,.., (⌊|NA|/|A|⌋ -1)}

Fig. 1. Our proposed Workflow

A. Data Collection

We use the UNSW-NB15 dataset [6, 7], collected by the IXIA PerfectStorm tool in the Cyber Range Lab, USA. It covers various network-related characteristics, such as source and destination IP addresses, protocol types, attack categories, origin and destination byte counts, source and destination port numbers, connection state tracking, duration, and numerous other pertinent variables. It comprises 49 distinct features, each providing specific network information. It is further enriched by a substantial volume of data instances, totaling 700,000 records that are a hybrid of the modern normal and abnormal network traffic in the form of packets.

B. Data Pre-processing and Cleaning

The data preparation process was initiated by selecting a dataset aligned with the network environment we aimed to model. During the initial examination, we noticed the presence of empty cells, particularly within the *attack_cat* feature. These blank cells corresponded to instances where no network attacks were detected. To ensure uniformity and facilitate our analysis, we replaced these empty cells with the label *no_attacks*. This *attack_cat* column plays a pivotal role in our study, as it directly reflects the collective impact of observed network behaviors and serves as the basis for label assignment. We use binary encoding to label the absence of an attack with *0*, while a label of *1* signifies the presence of a network attack.

C. Data Transformation

Converting categorical data into a format suitable for machine learning analysis is fundamental to ensuring that our models can effectively interpret and leverage categorical information. We applied one-hot encoding to the dataset, targeting the categorical columns *proto*, *state*, and *service*. One-hot encoding transforms these categorical features into a series of binary (*0* or *1*) columns, each representing a unique category. This transformation enables our machine learning algorithms to comprehend and process the categorical attributes to make accurate and meaningful predictions.

Further, we utilized *Min-Max* scaling, a fundamental preprocessing technique that guarantees that the numerical feature values are confined within a standardized range, typically from *0* to *1*, the columns that have been encoded using one-hot encoding. This prevents features with large numeric fields from influencing the analysis. We chose Min-Max scaling to provide uniformity in feature scales, which is crucial for machine learning algorithms like K-nearest neighbors (KNN).

This uniformity ensures all features contribute proportionally to our proposed K-nearest neighbors (KNN) DT model. We selectively applied Min-Max scaling to specific columns such as duration (*dur*), source bytes *sbyte*, destination bytes *dbyte*, and others. These columns were chosen based on their relevance to model creation. Columns with heterogeneous data or the potential to introduce bias (*proto*, *state*, and *service*) were excluded.

D. Data Balancing

Balancing of Data is a methodology employed in machine learning to rectify class imbalance within datasets. This imbalance manifests when certain classes within a classification problem exhibit fewer instances than others. Given the marked imbalance within the dataset employed for this research, it is deemed prudent to contemplate the creation of a balanced dataset derived from the original one. Upon scrutiny of the original dataset, the following observations were made concerning the label attribute consisting of anomaly (A) and not-anomaly (NA).

- Total number of 0's (Not anomaly) = 677,786
- Total number of 1's (Anomaly) = 22,215

We proposed a balancing method to balance the target variable, which has been compared with an existing method called the Synthetic Minority Over-sampling Technique (SMOTE). To obtain a balanced set, an equal number of Anomaly and Not Anomaly entries is required. In our case, since they are not equal; we took an equal number of Not Anomaly (677,786 - 655,571 = 22,215) to match the 22,215 entries of Anomaly.

$$NAS = \sum_{i=0}^{\lfloor \frac{|NA|}{|A|} \rfloor - 1} \{NA_{0+k}, NA_{1+k}, ..., NA_{|A|-1+k}\} + \epsilon \tag{1}$$

where $k = i * |A|$ and the value of ϵ has been ignored in this study.

E. Data Shuffling

Data shuffling is performed on each reduced 44,430 records of the balanced split dataset as follows:

$$BalancedDataset_i = suffle(A \cup NAS_i). \tag{2}$$

where each NAS_i is generated by Eq. 1. Further, each split $BalancedDataset_i$ is divided into an 80:20 ratio for the training and testing datasets, which gives appropriate training data for the small subsets of a UNSW-NB 15 dataset in our proposed method.

F. Model Fitting

The K-nearest Neighbour (k-NN) algorithm represents a non-parametric, supervised machine learning approach for tasks such as sample classification and regression. This sophisticated classifier evaluates the likeness between newly acquired data vectors and existing dataset entries. During its training phase, k-NN meticulously stores pertinent dataset information. Subsequently, when fresh data becomes available, the algorithm adeptly categorizes it into the most fitting category, closely aligning with the established dataset. The pivotal *k* parameter signifies the number of cases from the test or validation dataset that closely resemble a specific set of circumstances. The algorithm relies on the Euclidean distance measure to assess the similarity between data pairs. This method of proximity calculation proves instrumental in making informed categorization decisions based on the underlying dataset's intrinsic patterns.

The proposed KNN DT model learns from the training data, which consists of our selected numerical features, such as *dur*, *sbyte*, *dbyte*, and their corresponding labels. The training process enables our model to identify patterns and relationships within the data. The outcomes demonstrate the stability and accuracy with which our suggested model can identify network instances, hence augmenting the security of network systems.

IV. Model Evaluation

The proposed model is evaluated on the Synthetic Minority Over-sampling Technique (SMOTE) and compared with our proposed balanced datasets. SMOTE synthesizes the imbalanced data into a balanced one. As a result, it balanced the dataset after generating 22,215 anomalies into 677,786 to match with no anomaly entries. We found the accuracy of this synthesized dataset is 71.79%. However, our proposed balanced dataset method outperforms with 99.05% accuracy and 1.44% FPR.

Fig. 2 shows the values obtained for the confusion matrix after evaluating the proposed DT model on a split $BalancedDataset_i$, where i is a random dataset in $(0, 1, ..., (\lfloor |NA|/|A| \rfloor - 1)$. True positive (TP) represents the cases where our models accurately recognized and correctly identified anomalies. TP value of our proposed model on a balanced dataset is 4,435 out of 8,886 test dataset records, whereas correctly recognized non-anomalies (TN) is 4385. Type-I error occurs when our models mistakenly identify non-anomalies as anomalies. In this scenario, the Digital Twin model generates 49 false alarms, whereas incorrectly classifying (Type-II error) non-anomalous instances as anomalous are 15.

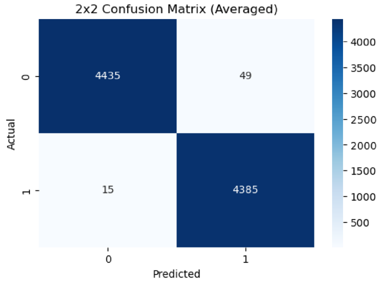
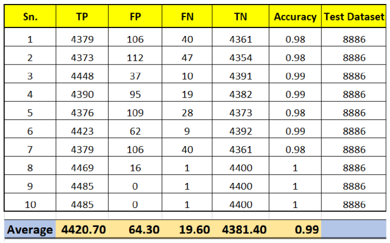


Fig. 2. Details of the confusion Matrix

Fig. 3 shows the average results of 10 datasets out of 30 $(= (\lfloor |NA|/|A| \rfloor - 1)$ to outline the evaluation of the proposed model. On average, the TP values of our proposed model on ten different balanced datasets are 4,420.70, whereas correctly recognized anomalies (TN) are 4381.40. The Digital Twin models, on average, generated 64.30 false alarms (False Positive) in this scenario, incorrectly classifying 19.60 anomalous instances as non-anomalous, whereas the proposed model missed an average of 15.08 anomalies (False Negative).



Sn.	TP	FP	FN	TN	Accuracy	Test Dataset
1	4379	106	40	4361	0.98	8886
2	4373	112	47	4354	0.98	8886
3	4448	37	10	4391	0.99	8886
4	4390	95	19	4382	0.99	8886
5	4376	109	28	4373	0.98	8886
6	4423	62	9	4392	0.99	8886
7	4379	106	40	4361	0.98	8886
8	4469	16	1	4400	1	8886
9	4485	0	1	4400	1	8886
10	4485	0	1	4400	1	8886
Average	**4420.70**	**64.30**	**19.60**	**4381.40**	**0.99**	

Fig. 3. Evaluation Report of the proposed DT model on ten different balanced datasets

Fig. 4 shows the values of *precision*, *recall*, and *F1-score* of the non-anomaly situation. Precision is the ratio of correctly predicted positive values among a model's total number of positive predictions (TP / (TP + FP)). This metric quantifies the proportion of correctly predicted positive instances among all positive predictions. In our study, the average is 1.00 for no anomaly detected and 0.99 for abnormality detected precisely. A high precision indicates that the models have a low rate of false positives.

Recall, also called sensitivity, represents the ratio of true positive predictions to the total number of positive instances in the dataset (TP / (TP + FN)). It quantifies the accuracy in identifying actual positive instances as positive. In our analysis, the average equals 0.99 for both anomaly and no anomaly detected. A high recall value signifies the models' proficiency in identifying positive instances. Finally, the F1-score is the harmonic mean of precision and recall to visualize their combined effect on the proposed model. It balances precision and recall and is helpful to find a balance between false positives and false negatives. The report shows 0.99 on average for both anomalies and no anomaly detection.

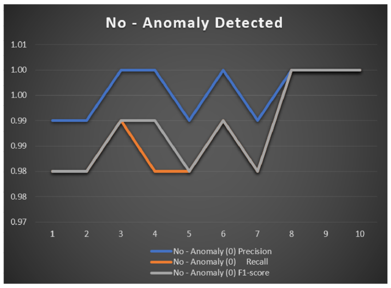


Fig. 4. Graphical view of No-Anomaly Detection

Moreover, Fig 4 reflects how the models meticulously captured normal situations with good precision growth. This is one of the benefits of using a balanced dataset in model creation. KNN exhibits excellent performance when given a balanced input dataset to train since it relies on the distance between vector points to build its analogy. Fig. 5 shows the range of precision from 0.97 to 1. We observe a usual situation of KNN models struggling to identify anomalies precisely. Since this research's primary objective is the proactive identification of anomalies within a data stream using the DT models, precision from the onset might face some challenges.

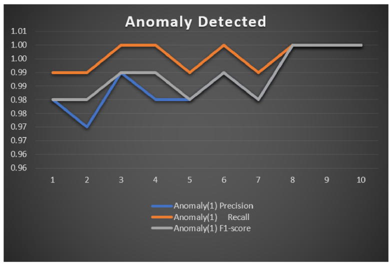


Fig. 5. Anomaly Detection in the proposed DT model

A. Evaluation using AUROC Curve

Fig. 6 describes the Receiver Operating Characteristics. An Area Under the Receiver Operating Characteristic (AUROC) score of 1.00 indicates that the Digital Twin (DT) model's performance in distinguishing between anomalies (attacks) and non-anomalies (non-attacks) is perfect. It means that the model has achieved a flawless balance between true positives and false positives, as well as true negatives and false negatives.

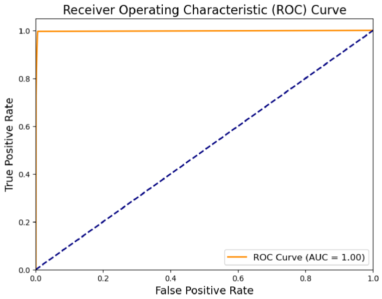


Fig. 6. ROC curve for the proposed DT model

The previously presented confusion matrix and classification report demonstrates the DT model's remarkable ability to accurately differentiate between anomalies and non-anomalies. The precision, recall, and F1-score values for both classes (0 and 1) were notably high, signifying a low occurrence of false positives and false negatives. Consequently, the ROC curve, which visualizes the trade-off between true positive rate and false positive rate, demonstrates a curve that hugs the top-left corner of the plot, ultimately leading to an AUROC score of 1.00.

Further, the Precision-Recall curve is a crucial evaluation tool for our Digital Twin (DT) model in the context of anomaly detection. Fig. 7 shows the Precision-recall of the Digital Twin model. This diagram illustrates the balance between precision and recall across different decision thresholds, offering valuable insights into the model's performance. Precision, in this context, measures the accuracy of the DT model's positive predictions, specifically the ratio of true positive predictions (correctly identified anomalies) to all instances classified as anomalies. Recall quantifies the DT model's ability to identify all actual anomalies correctly, represented as the ratio of true positive predictions to all actual anomalies.

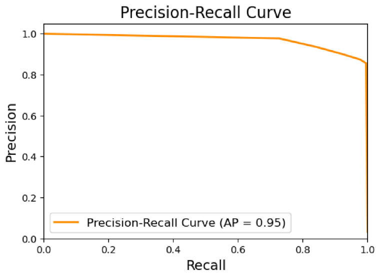


Fig. 7. The Precision-Recall Curve of the proposed DT model

The Precision-Recall curve is essential because it allows us to assess our model's performance under different levels of anomaly detection strictness. It helps us balance minimizing false alarms (improving precision) and ensuring that actual anomalies are not missed (improving recall). The Area Under the Precision-Recall Curve (AP) is a singular metric that encapsulates the model's overall performance across a range of thresholds. In our case, an AP of 0.95 is highly commendable. It signifies that the DT model consistently achieves high precision while maintaining strong recall, even when adjusting the decision threshold.

In practical terms, a high AP score like 0.95 means that the DT model is exceptionally effective at identifying network anomalies with a low rate of false positives. It is precious in network security, as it indicates that the model can accurately pinpoint security threats while minimizing unnecessary alerts. We identified anomalies using key terminologies such as *Receiver*, *Destination*, and *Source*.

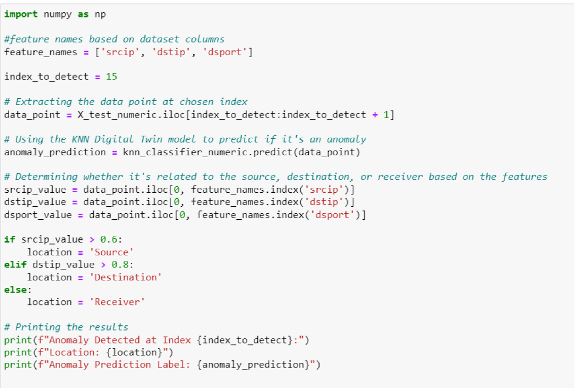
- *Receiver:* The designation is applied when neither the *Source* nor the *Destination* can be conclusively identified as the primary contributor to an anomaly. When neither the *srcip* value (source IP address) nor the *dstip* value (destination IP address) surpasses predetermined thresholds (0.6 and 0.8, respectively), the location is marked as *Receiver*. It indicates that the anomaly is not explicitly linked to the data source or destination but is considered an anomaly at an intermediate point within our dataset or network.
- *Source:* It is invoked when the *srcip* value (source IP address) exceeds a specific threshold (e.g., 0.6). It suggests that the anomaly is related to the source of the data point, typically where data originates. If *srcip_value* exceeds the defined threshold, the location is attributed as *Source*. It signifies that the anomaly arises from the source IP address, representing its association with the data's point of origin.
- *Destination:* It is employed when the *dstip* value (destination IP address) surpasses a certain threshold (e.g., 0.8). It indicates that the anomaly is associated with the destination of the data point, where data is received or directed. If *dstip_value* surpasses the predefined threshold, the location is denoted as *Destination*. It implies that the anomaly is linked to the destination IP address, indicating its association with the data's intended endpoint or destination.

These location attributions serve as critical indicators, aiding us in understanding the origins of anomalies—whether they manifest at the source, destination, or intermediary points within the dataset. They augment the depth and specificity of our research, facilitating a more thorough examination of the spatial dimensions of unusual behavior within our dataset.

Fig. 8 depicts the code snippet used for the exercise. The proposed DT model was tested over sample data to identify anomalies. When it was limited to index 15, it could detect anomalies at that index and label it as *1*, clarifying the details of the classification report where the anomaly is class 1 (Fig. 9).

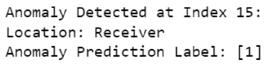
V. Conclusion

The KNN-based Digital Twin model showcases remarkable performance, accurately discerning non-anomalies while preserving a favorable trade-off between precision and recall



```
import numpy as np

#feature names based on dataset columns
feature_names = ['srcip', 'dstip', 'dsport']

index_to_detect = 15

# Extracting the data point at chosen index
data_point = X_test_numeric.iloc[index_to_detect:index_to_detect + 1]

# Using the KNN Digital Twin model to predict if it's an anomaly
anomaly_prediction = knn_classifier_numeric.predict(data_point)

# Determining whether it's related to the source, destination, or receiver based on the features
srcip_value = data_point.iloc[0, feature_names.index('srcip')]
dstip_value = data_point.iloc[0, feature_names.index('dstip')]
dsport_value = data_point.iloc[0, feature_names.index('dsport')]

if srcip_value > 0.6:
    location = 'Source'
elif dstip_value > 0.8:
    location = 'Destination'
else:
    location = 'Receiver'

# Printing the results
print(f"Anomaly Detected at Index {index_to_detect}:")
print(f"Location: {location}")
print(f"Anomaly Prediction Label: {anomaly_prediction}")
```

Fig. 8. Details of the code used to detect anomaly at index 15



```
Anomaly Detected at Index 15:
Location: Receiver
Anomaly Prediction Label: [1]
```

Fig. 9. Results of the anomaly detected by the proposed DT model

for anomalies. The model's high accuracy and robust F1 scores underscore its reliability in anomaly detection. Furthermore, the AUROC score of 1.00 signifies the proposed model's exceptional ability to make precise predictions with minimal declassifications, achieving an optimal equilibrium between sensitivity and specificity. It emphasizes the model's unwavering capacity to differentiate normal network behavior from potential security threats. Moreover, it has an accuracy of 99.05% with its impressive AP score of 0.95, which is comparatively higher than the accuracy of existing works (that is 86% [17] and 98.97% [18]). The Precision-Recall curve underscores the model's effectiveness in achieving precision and recall. It positions it as a potent asset for network anomaly detection, significantly enhancing overall network security. Our next goal is to implement the proposed ML-based DT model into a live network to predict and detect anomalies in real-time.

References

[1] Holmes, M. Papathanasaki, L. Maglaras, M. A. Ferrag, S. Nepal and H. Janicke, ”Digital Twins and Cyber Security – solution or challenge?,” 2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), Preveza, Greece, 2021, pp. 1-8, doi: 10.1109/SEEDA-CECNSM53056.2021.9566277.

[2] Maschler, D. Braun, N. Jazdi, and M. Weyrich, “Transfer Learning as an Enabler of the Intelligent Digital Twin,” no. November, 2020. [Online]. Available: http://arxiv.org/abs/2012.01913.

[3] Ashtari Talkhestani B, Tobias, J., Lindemann, B., Nada, S., Nasser, J.,Wolfgang, S., and Michael, W., An architecture of an Intelligent Digital Twin in a Cyber-Physical Production System. at - Automatisierungstechnik 2019; 9:762–82.

[4] Kerpicci, M., Ozkan, H., Kozat, S.S., 2021. ”Online anomaly detection with bandwidth optimized hierarchical kernel density estimators”. IEEE Trans. Neural Netw. Learn. Syst. 32 (9), 4253–4266.

[5] International Atomic Energy Agency, 2019. PCTRAN Generic Pressurized Water Reactor Simulator Exercise Handbook, IAEA-TCS-68. IAEA, Vienna.

[6] https://cloudstor.aarnet.edu.au/plus/index.php/s/2DhnLGDdEECo4ys, accessed on September 2023.

[7] Moustafa, Nour, and Jill Slay. ”UNSW-NB15: a comprehensive dataset for network intrusion detection systems (UNSW-NB15 network dataset).” Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015.

[8] Leukel, J., González, J., Riekert, M., 2021. Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review. J. Manuf. Syst. 61, 87–96.

[9] Xia, M., Shao, H., Williams, D., Lu, S., Shu, L., de Silva, C.W., 2021. Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning. Reliab. Eng. Syst. Saf. 215, 107938.

[10] Barbado, A., Corcho, Ó., Benjamins, R., 2022. Rule extraction in unsupervised anomaly detection for model explainability: application to OneClass SVM. Expert Syst. Appl. 189, 116100.

[11] Almasan,P, etal (2022).Digitaltwinnetwork Opportunitiesandchallenges.arXivpreprintarXiv 2201.01144.

[12] Eckhart, M., Ekelhart, A., 2019. Digital twins for cyber-physical systems security: State of the art and outlook. In: Biffl, S., Eckhart, M., Lüder, A., Weippl, E. (Eds.), Security and Quality in Cyber-Physical Systems Engineering: With Forewords by Robert M. Lee and Tom Gilb. Springer International Publishing, Cham, pp. 383–412. http://dx.doi.org/10.1007/978-3-030-25312-7_14.

[13] Lee, R.M., Assante, M., Conway, T., 2017. Crashoverride: Analysis of the Threat to Electric Grid Operations. Dragos Inc., [Online]. Available: https://www.dragos.com/wp-content/uploads/CrashOverride-01.pdf.

[14] M. Eckhart and A. Ekelhart, “Towards security-aware virtual environments for digital twins,” in Proceedings of the 4th ACM Workshop onCyber-Physical System Security, ser. CPSS ’18. New York, NY, USA: ACM, 2018, pp. 61–72.

[15] M. Eckhart and A. Ekelhart, “Securing cyber-physical systems through digital twins,” ERCIM News, vol. 2018, no. 115, 2018. [Online]. Available: https://ercim-news.ercim.eu/en115/special/ 2101-securing-cyber-physical-systems-through-digital-twins.

[16] M. Eckhart and A. Ekelhart, “A specification-based state replication approach for digital twins,” in Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy, ser. CPS-SPC ’18. New York, NY, USA: ACM, 2018, pp. 36–47.

[17] Farhat, M. H., Chiementin, X., Chaari, F., Bolaers, F. and Haddar, M. (2021). Digital twin-driven machine learning: ball bearings fault severity classification. Measurement Science and Technology, 32(4), 044006.

[18] Moustafa N, Turnbull B, Choo KK. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. IEEE Internet of Things Journal. 2018 Sep 23; 6(3):4815-30.