Anomaly Detection in Software-Defined Networks

Subjects: Computer Science, Information Systems

Contributor: Grzegorz Rzym , Amadeusz Masny , Piotr Chołda

One solution enabling the implementation of modern methods for managing and monitoring telecommunication networks is the concept of software-defined networks (SDN). It introduces a centralized architecture for managing computer networks that is fully programmable. With the increasing availability of computational power, contemporary machine learning has undergone a paradigm shift, placing a heightened emphasis on deep learning methodologies. The pervasive automation of various processes necessitates a critical re-evaluation of contemporary network implementations, specifically concerning security protocols and the imperative need for swift, precise responses to system failures.

software-defined network (SDN) anomaly detection deep learning machine learning algorithms

1. Introduction

Traditional methods of managing computer networks and monitoring these networks rely on data collection protocols such as SNMP (Simple Network Management Protocol), NetFlow, IPFIX, or NETCONF (Network Configuration Protocol) [1]. Many companies create their own tools based on these protocols, allowing insight into the state of their computer network and responding to failures. With the impressive increase in demand for network services, requirements for individual components and monitoring applications have also increased.

The amount of data generated by network devices, including control, statistical, and user data, is growing incredibly fast [2][3][4][5]. Traditional human-involved analysis methods require qualified human resources and working time, a trend increasingly displaced by automated solutions in today's era of computerization.

One solution enabling the implementation of modern methods for managing and monitoring telecommunication networks is the concept of software-defined networks (SDN) [6]. It introduces a centralized architecture for managing computer networks that is fully programmable. However, the analytical aspect remains, involving continuous network monitoring and anomaly detection. Collecting data in one central location has business benefits, facilitating the coordination and correlation of collected results. Nevertheless, searching such a database can pose a significant computational challenge known as "big data analytics" [7]. Tools facilitating the implementation, deployment, and utilization of massive datasets come to the rescue. An example is the open source solution proposed by Cisco—the Platform for Network Data Analytics (PNDA) [8]. PNDA allows for the

collection and reading of data, along with real-time analysis. Considering the enormous amount of data for analysis, this is where the possibility of harnessing artificial intelligence (AI) and machine learning (ML) comes into play.

Continuous data collection is an ideal solution for monitoring systems. However, it can be very disadvantageous in terms of device load and network throughput, whether it be in the transmission or management network. Therefore, dynamically changing data collection intervals based on the network's state (e.g., increasing polling frequency after anomaly detection) seems reasonable.

2. Machine Learning Applied to SDNs

Machine learning applied to SDNs represents a cutting-edge approach to enhancing network performance, efficiency, and security. In the realm of SDNs, machine learning algorithms analyze vast amounts of network data, adapting and optimizing the network dynamically. These algorithms can predict traffic patterns, identify anomalies, and automate network management tasks. For an in-depth exploration of the applications of machine learning in software-defined networks, refer to the comprehensive survey presented in [9].

The exploration of the deployment of 6G networks is anticipated to usher in transformative enhancements to network architectures, with a specific emphasis on the integration of AI technologies. Reference [10] delves into the novel concept of knowledge-defined networking (KDN), wherein network intelligence is concentrated in the knowledge plane, achieved through a fusion of SDN, network telemetry, and ML algorithms. Notably, this research underscores the utility of programming protocol-independent packet processors (P4), a technology facilitating SDN networks, and underscores the significance of in-band network telemetry (INT) for furnishing real-time network insights. Furthermore, it establishes a link between P4-SDN network architecture and reinforcement learning (RL), illustrating how network components and established techniques can be aligned with RL principles. The research also delves into the potential of AI-driven network orchestration and expounds upon the conceptualization of networks as AI-based systems. While the research outlined in [10] offers a comprehensive framework for innovative mobile networks, it suggests the utilization of INT for latency measurement in networks. However, it is crucial to note that this framework remains in the proposal stage, lacking the implementation details or simulation results presented in this research.

The research presented in [11] proposes a strategy to tackle the challenge of unpredictable topology states in Flying Ad Hoc Networks (FANET). The researchers of this research deployed an AI algorithm capable of discerning patterns in unmanned aerial vehicle (UAV) mobility, anticipating potential disconnections and proactively initiating rerouting or forwarding algorithms. The research introduces a case of a software-defined FANET that offers wireless INT to an AI-equipped edge node situated at the ground station. It elaborates on the design of the subsystems housing the AI process and illustrates how a machine learning model can identify critical network situations without reliance on intricate neural networks.

In [12], the researchers introduced integration of SDN with INT and deep reinforcement learning (DRL) to autonomously manage and enhance network performance. A QoS-routing use case and preliminary experimental evidence are presented to demonstrate the feasibility of the proposed paradigm. Additionally, some important challenges that need to be addressed are discussed. The researchers advocate that addressing such challenges requires a truly interdisciplinary effort between the research fields of artificial intelligence, network science, and computer networks.

The researchers of [13] assessed an ML-based soft-failure localization framework in scenarios involving partial telemetry. The framework, based on an artificial neural network (ANN), is trained using optical signal and noise power models simulating network telemetry across all potential failure scenarios. The ML-based framework demonstrates exceptional performance in partial telemetry scenarios, effectively interpolating missing data. The research also demonstrates that ANN training is expedited by principal component analysis and can be conducted using cloud-based services. Additionally, the researchers emulated the evaluated ML-based framework in a software-defined networking-based setup using the gRPC Network Management Interface protocol for streaming telemetry.

In the research conducted by Faheem et al. [14], the researchers explored a spectrum of machine learning techniques dedicated to estimating the resource requirements of intricate network entities, particularly Virtual Network Functions (VNFs) within a software-defined networking environment. Their focus primarily centered on deciphering the resource demands of VNFs, notably the central processing unit (CPU) consumption during the processing of input traffic. The experiments conducted in their research not only underscored the ML models' aptitude for learning the intricate behaviors of VNFs but also showcased their efficacy in accurately modeling the resource requirements. The findings put forth by the researchers suggest that ML techniques can serve as highly effective tools for modeling the resource needs of diverse VNFs, providing valuable insights into optimizing resource allocation and enhancing the overall efficiency of network environments.

The research conducted by Alshahrani et al. [15] provides a comprehensive examination of the complexities introduced by the smart city initiative, characterized by a myriad of specifications and a diverse user base with distinct requirements. To address the challenges arising from this dynamic environment, the researchers propose an innovative system that integrates SDN security controllers and ML models with optimization techniques. This strategic combination aims to effectively mitigate the impact of prevalent Distributed Denial of Service (DDoS) attacks on smart cities. The proposed approach is built upon an SDN infrastructure supported by security controllers, constituting a proactive line of defense against potential threats. Additionally, the detection mechanism embedded in the ML model, optimized for enhanced performance, plays a pivotal role in identifying and neutralizing common DDoS attacks within smart city networks. This dual-layered security strategy demonstrates the proposed system's capability to protect against evolving cybersecurity threats. Furthermore, Alshahrani et al. advocate for the implementation of binary classification as a crucial component of their proposed system. The adoption of this classification method not only enhances the efficiency of attack detection but also results in a commendable level of accuracy.

Given the increasing prevalence of Internet of Things (IoT) devices connected to the Internet, the annual rise in IoT-based attacks has prompted a need for more effective solutions. Existing approaches may struggle to sufficiently mitigate these attacks, especially in network environments supporting both traditional and IoT protocols, and utilizing a centralized architecture like SDN. The research in [16] introduces a long short-term memory (LSTM)-based approach for detecting network attacks within IoT networks using an SDN-supported intrusion detection system. The performance of the machine learning and deep learning model is evaluated across two SDNIoT-focused datasets. Additionally, an LSTM-based architecture is proposed for the effective multiclass classification of network attacks in IoT networks. The evaluation demonstrates the model's effectiveness in identifying and classifying attacks, achieving a high level of accuracy. This research also employs various visualization methods to comprehend dataset characteristics and visualize embedding features.

Various models have been employed in the literature to identify anomalies. For instance, the paper presented in [17] proposes the Two-Step Graph Convolutional Neural Network (TS-GCN) framework. This framework, incorporating resampling techniques and adopting a streamlined architecture, establishes itself as the benchmark for addressing the identified problem. When applied to a specific satellite model, TS-GCN demonstrates notable success in state recognition and prediction accuracy. In comparison to established models, TS-GCN showcases considerable enhancements in state recognition accuracy. The conclusion suggests that TS-GCN, with its streamlined architecture and applicability for on-orbit deployment, holds promise for improved assessment and anomaly detection in satellite systems.

The researchers of [18] explore the use of Markov models for anomaly detection in the Healthcare Internet of Things (HIoT). The proposed method leverages the simplicity, interpretability, and a well-developed theory of Markov models to enhance cybersecurity in HIoT. By evaluating the method using the ToN_IoT dataset, the research aimed to address security concerns and contribute to safeguarding patients' wellbeing in healthcare services.

In [19], the researchers delve into the application of undirected probabilistic graphical models, specifically the residual Gauss–Markov random field, for characterizing cloud telemetry. Acknowledging the complexities of cloud systems, the research proposes a unique data model and outlines an efficient estimation procedure. The primary focus is on anomaly detection and localization, demonstrated through experiments in synthetic and small-scale software system environments. The research underscores the computational attractiveness of fitting the model and addresses practical considerations in cloud system structure. However, it also highlights the challenge of validating anomaly detection techniques under the constraints of real-world production cloud systems.

The research presented in [20] addresses the challenges of anomaly detection in satellite telemetry data, proposing a novel model based on Bayesian deep learning. The model leverages Monte Carlo Dropout on a long short-term memory network (LSTM) and establishes the Bayesian LSTM for effective anomaly detection without domain knowledge. The research introduces the concept of uncertainty measures, including Monte Carlo Sampling Variance, Prediction Entropy, and Mutual Information, to enhance anomaly detection capabilities. Additionally, the research explores these uncertainties further and employs a variational auto-encoder (VAE) to re-evaluate high-uncertainty samples, improving the model's robustness on imbalanced datasets. The experimental results

demonstrate the effectiveness of the proposed model, showcasing its superior performance over traditional neural networks and other Bayesian neural networks in handling imbalanced datasets.

References

1. Fernandes, G.; Rodrigues, J.J.; Carvalho, L.F.; Al-Muhtadi, J.F.; Proença, M.L. A Comprehensive Survey on Network Anomaly Detection. Telecommun. Syst. 2019, 70, 447–489.
2. Cisco Annual Internet Report (2018–2023) White Paper; Technical Report; Cisco: San Jose, CA, USA, 2023.
3. 2022 Global Networking Trends Report; Technical Report; Cisco: San Jose, CA, USA, 2022.
4. 2023 Global Internet Phenomena Report; Technical Report, Sandvine Intelligent Broadband Networks; Sandvine Inc.: Waterloo, ON, Canada, 2022.
5. Ericsson Mobility Report; Technical Report; Ericsson: Stockholm, Sweden, 2022.
6. Karakus, M.; Durresi, A. A Survey: Control Plane Scalability Issues and Approaches in Software-Defined Networking (SDN). Comput. Netw. 2017, 112, 279–293.
7. Cui, L.; Yu, F.R.; Yan, Q. When big data meets software-defined networking: SDN for big data and big data for SDN. IEEE Netw. 2016, 30, 58–65.
8. Project PNDA Web Page. Available online: https://pnda.io (accessed on 15 December 2023).
9. Pathak, Y.; Prashanth, P.V.N.; Tiwari, A. AI Meets SDN: A Survey of Artificial Intelligent Techniques Applied to Software-Defined Networks. In 6G Enabled Fog Computing in IoT: Applications and Opportunities; Kumar, M., Gill, S.S., Samriya, J.K., Uhlig, S., Eds.; Springer Nature Switzerland: Cham, Switzerland, 2023; pp. 395–412.
10. Zeman, D.; Zelinka, I.; Voznak, M. A Reinforcement Learning Framework for Knowledge-Defined Networking. In Proceedings of the 2023 15th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Ghent, Belgium, 30 October–1 November 2023; pp. 152–156.
11. Uomo, D.; Sgambelluri, A.; Castoldi, P.; De Paoli, E.; Paolucci, F.; Cugini, F. Failure Prediction in Software Defined Flying Ad-Hoc Network. In Proceedings of the Twenty-Fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing, New York, NY, USA, 23–26 October 2023; MobiHoc '23. pp. 355–357.
12. Yao, H.; Mai, T.; Xu, X.; Zhang, P.; Li, M.; Liu, Y. NetworkAI: An Intelligent Network Architecture for Self-Learning Control Strategies in Software Defined Networks. IEEE Internet Things J. 2018, 5, 4319–4327.

13. Mayer, K.S.; Soares, J.A.; Pinto, R.P.; Rothenberg, C.E.; Arantes, D.S.; Mello, D.A.A. Machine-learning-based soft-failure localization with partial software-defined networking telemetry. J. Opt. Commun. Netw. 2021, 13, E122–E131.
14. Faheem, S.M.; Babar, M.I.; Khalil, R.A.; Saeed, N. Performance Analysis of Selected Machine Learning Techniques for Estimating Resource Requirements of Virtual Network Functions (VNFs) in Software Defined Networks. Appl. Sci. 2022, 12, 4576.
15. Alshahrani, M.M. A Secure and Intelligent Software-Defined Networking Framework for Future Smart Cities to Prevent DDoS Attack. Appl. Sci. 2023, 13, 9822.
16. Chaganti, R.; Suliman, W.; Ravi, V.; Dua, A. Deep Learning Approach for SDN-Enabled Intrusion Detection System in IoT Networks. Information 2023, 14, 41.
17. Liu, S.; Qiu, S.; Li, H.; Liu, M. Real-Time Telemetry-Based Recognition and Prediction of Satellite State Using TS-GCN Network. Electronics 2023, 12, 4824.
18. Huang, H.C.; Liu, I.H.; Lee, M.H.; Li, J.S. Anomaly Detection on Network Traffic for the Healthcare Internet of Things. Eng. Proc. 2023, 55, 3.
19. Landolfi, N.C.; O'Neill, D.C.; Lall, S. Cloud Telemetry Modeling via Residual Gauss-Markov Random Fields. In Proceedings of the 2021 24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN), Paris, France, 1–4 March 2021; pp. 49–56.
20. Chen, J.; Pi, D.; Wu, Z.; Zhao, X.; Pan, Y.; Zhang, Q. Imbalanced satellite telemetry data anomaly detection model based on Bayesian LSTM. Acta Astronaut. 2021, 180, 232–242.