METINĖ ATASKAITINĖ INFORMATIKOS KRYPTIES DOKTORANTŲ KONFERENCIJA 2018 M. SPALIO 24 D.

ATASKAITA

DOKTORANTAS VIKTORAS BULAVAS
INFORMATIKA (09P)

VADOVAS: PROF. HABIL. DR. GINTAUTAS DZEMYDA

KONSULTANTAS: DR. VIRGINIJUS MARCINKEVIČIUS

DOKTORANTŪROS LAIKOTARPIS 2017 M. - 2021 M.

DMSTI-DS-09P-18-1

Tyrimo objektas, tikslai ir planuojami gauti rezultatai

- Preliminari disertacijos tema ir tyrimo objektas:
  - **Mašininio mokymo metodų taikymas ankstyvajam kibernetinių incidentų aptikimui**
- Tyrimo tikslai:
  - Gauti naujos informacijos apie tinkamus ankstyvojo anomalijų aptikimui mašininio mokymosi metodus
- Planuojami gauti rezultatai:
  - Panaudoti parinktus metodus, siekiant prognozuoti bei valdyti ankstyvąjį kibernetinių incidentų etapą

Ataskaitinių metų darbo planas

- Mokslinių tyrimų disertacijos tema apžvalga ir analizė:
  - Disertacijos tyrimo objekto detalizavimas.
  - Mašininio mokymosi metodų taikymo kibernetinės saugos ankstyvo įspėjimo problemoms spręsti metodų apžvalga.
- Publikacijos parengimas konferencijos medžiagoje Lietuvoje.
- 2 egzaminai (16 kreditų).
- Bendrųjų gebėjimų mokymai.

Atlikta: Egzaminai

- 2018 m. birželio 4 d. išlaikytas egzaminas „Duomenų analizės strategijos ir sprendimų priėmimas“, 7 kreditai, įvertinimas – puikiai (10)
- 2018 m. birželio 28 d. išlaikytas egzaminas „Informatikos ir informatikos inžinerijos tyrimo metodai ir metodika“, 9 kreditai, įvertinimas – gerai (8)

Atlikta: DMSTI pristatymai

- DMSTI 2017-10-23 Naujų doktorantų prisistatymas
- DMSTI 2018-03-05, Ankstyvojo kibernetinių incidentų aptikimo metodai, pristatymas

Atlikta: konferencijos (1)

- 2017 m. gruodžio 1 d. 9-oje konferencijoje "Duomenų analizės metodai programų sistemoms", Druskininkai, pristatytas stendinis pranešimas „An investigation of Early Cyber Threat Detection using Ensembles of Machine Learning Methods“

An investigation of Early Cyber Threat Detection using Ensembles of Machine Learning Methods

V. Bulavas, G.Dzemyda, V. Marcinkevičius
Vilnius University Faculty of Mathematics and informatics
Institute of Data Science and Digital Technologies
Vilnius University

Abstract

According to PwC's Global economic crime survey, cybercrime has overall evolved into the second place after an asset misappropriation. According to Lithuanian National Cyber Security Centre Annual report for 2016, scanning of surveilled network devices since 2015 has increased fivefold. Lithuanian academic network LITNET is no different, observing persistent multiple step intrusion activities. As nowadays it is impossible to detect and mitigate all threats manually, automatic tools are used on a 24/7 basis. The techniques utilized by current network intrusion detection appliances in use fall into three main categories: anomaly detection, misuse detection and hybrid. Misuse detection systems use signatures that describe already known attacks and require regular ruleset update. Machine learning based anomaly detection requires supervision and specialist review due to currently still high false positive rate of detecting previously unseen system behaviors. With an increasing frequency of cyber-attack, reviews take more and more time of cyber security specialists, which is a challenge. This indicates highly demanded area for research aiming to increase threat detection accuracy and training speed. Until very recently there was little published research about successful early threat detection models. Other authors proposed an ensemble of Machine Learning models as a probable way for solving abovementioned early detection problems. Therefore, in this work, authors perform investigation of selected method ensembles and present results of comparison.

Keywords

Network security; intrusion detection; early warning; anomaly detection; machine learning; ensemble learning; neuron capsules

Objectives

Response to cyber crime, with an increasing speed becoming everyday hassle to the society and enterprises, requires multi-layered Intrusion Detection Systems (IDS) [Figure 1].

Figure 1. Layers of a Next-Generation IDS [1].

In this research we pay attention to Early Warning layer of network based intrusion detection. The basic style attack types, occurring in cyber space require different types of Machine Learning (ML) algorithm ensembles to reduce the false alert rate (FAR), which varies from 0% to 5% depending on the ML method used and the type of attack. Meanwhile scanning of surveilled network devices is increasing, new threats are addressing diverse multi-layer attack vectors and require intensive use of behavior-based detection techniques. Current objective of research is to investigate the new neuron capsule method for possible application in Intrusion Detection Systems and possibly increase the speed of threat identification.

Issues of Flow-based Detection

The router or switch has the ability to collect IP network traffic as it enters and exits the interface. Flow monitoring has become a prevalent method for monitoring traffic in high-speed networks. A network flow is predominantly defined as a unidirectional sequence of packets that share the exact same packet attributes: ingress interface, source IP address, destination IP address, IP protocol, source port, destination port, and IP type of service. The NetFlow protocol itself has been superseded by Internet Protocol Flow Information eXport (IPFIX), therefore ML training data sets created earlier are becoming obsolete and are far from being universal. Even though NetFlows may be still the most frequent due to Cisco's popularity in the networking industry, other network equipment vendors provide similar network flow monitoring technology, which implies, that flows training has to be tailored for specific equipment on site.

Ensemble Methods for IDS

Ensemble learning methods train combinations of base models traditionally used in supervised learning [Figure 2].

Figure 2. Ensemble of ML Methods.

Machine learning approaches used in intrusion detection include Decision Trees, Inductive Learning, Naive Bayes, Random Forest, Artificial Neural Networks, Fuzzy Systems, Evolutionary Computation, Artificial Immune Systems, Hidden Markov, Sequential Pattern Mining, Swarm Intelligence [2], [3] and other.

Ensembles of ML methods were demonstrated to be an efficient way of improving predictive accuracy and/or decomposing a complex, difficult learning problem into some easier tasks [4]. According to the "no free lunch" theorem, there is not a single classifier that is appropriate for all the tasks, since each algorithm has its own domain of competence. Therefore we need a pool of classifiers to solve a given problem.

However there are several known ensemble learning issues such as the number and types of base models to use, the combining method to use, and how to maintain diversity among the base models. Current IDS require immense amount of data to learn, and data from one source (or location) is not enough. Experts are concerned with a need of constant retraining for IDS and refeeding same data into different models of an Ensemble.

Neuron Capsules

Facing the fact, that network data flows are coming as a stream of virtually never repeating data, the relatively new concept of neuron capsules is expected to help overcoming the limitation of a need of constant retraining. S. Sabour, N. Frosst and G.E. Hinton [5] demonstrated that a discriminatively trained, multi-layer capsule system achieves superior performance, reducing the number of test errors by 45% compared to the previously used ML methods applied in image recognition area.

A capsule is a group of neurons whose outputs represent different properties of the same entity. Active capsules at one level make predictions for the instantiation parameters of higher level capsules. When multiple predictions agree, a higher level capsule becomes active. To achieve these results an iterative routing by agreement mechanism is applied: a lower level capsule prefers to send its output to higher level capsules whose activity vectors have a big scalar product with the prediction coming from the lower level capsule [5].

The Capsule Network is expected to be capable of extracting more understanding from a given amount of data than single Ensemble.

Future Work

This research allows us to substantiate, that if using neuron capsules helped solving classification problem for image data, it is highly probable, that it will be effective with classification of network traffic data.

Therefore the objective of our future research is to investigate the Neuron Capsule method application for early warning and intrusion detection tasks.

For that we need to build a test environment, train the system and calculate indicators for comparison of classical Ensemble methods and the new Capsule Networks.

Literature

1. Koch, R. (2011). *Towards Next-Generation Intrusion Detection*. In T. W. (Eds. C. Czosseck, E. Tyugu (Ed.), 2011 3rd International Conference on Cyber Conflict (pp. 151–168). Tallinn, Estonia: CCD COE Publications.
2. F. Gharibian and A. Ghorbani (2007). *Comparative study of supervised machine learning techniques for intrusion detection*, in Proc. 5th Annu. Conf. Commun. Netw. Serv. Res., pp. 350–358.
3. Buczak, A., & Guven, E. (2015). *A survey of data mining and machine learning methods for cyber security intrusion detection*. IEEE Communications Surveys & Tutorials, pp. (1153-1176)
4. Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., & Woźniak, M. (2017). *Ensemble learning for data stream analysis: A survey*. Information Fusion, 37, 132–156.
5. Sabour, S., Frosst, N., & Hinton, G. E. (2017). *Dynamic Routing Between Capsules*. Retrieved from http://arxiv.org/abs/1710.09829

Atlikta: konferencijos (2)

- 2018 m. spalio 12 dieną konferencijoje „The 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University“, Rygoje, pristatytas pranešimas „Investigation of network intrusion detection using data visualization methods“.

1862
RĪGAS TEHNISKĀ UNIVERSITĀTE

IEEE
Advancing Technology for Humanity

CERTIFICATE OF PARTICIPATION

issued to

Viktoras Bulavas,

Vilnius University
Lithuania

in recognition of his/her participation in

The 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS'2018)

held at Riga Technical University, Latvia on October 10-12, 2018

Presentation title: **Investigation of Network Intrusion Detection Using Data Visualization Methods**
by Viktoras Bulavas

Prof. Dr. Jānis Grabis
ITMS'2018 General Chair
Riga Technical University, Latvia

IEEE Conference record number: 45466

Bendrųjų gebėjimų mokymai doktorantams
3,5 ECTS kredito

- ,,Mokslinių rezultatų publikavimas" - 2017 m. lapkričio 16 d., 5 akad. val., 0,25 ECTS kredito;
- ,,Intelektinės nuosavybės apsauga" - 2017 m. gruodžio 14 d., 5 akad. val., 0,25 ECTS kredito;
- ,,Vertės pasiūlymas. Kas tai yra ir kaip jį sukurti?" - 2018 m. vasario 8 d., 5 akad. val., 0,25 ECTS kredito;
- ,,[vadas į R" - 2018 m. kovo 15 d. ir 2018 m. kovo 22 d., 32 akad. val., 1,25 ECTS kredito;
- ,,Darbas su LaTeX" - 2018 m. balandžio 5 d. ir 2018 m. balandžio 12 d., 32 akad. val. 1,25 ECTS kredito;
- ,,Mokslo projektai - finansinės priemonės ir paraiškų rengimas" - 2018 m. balandžio 9 d., 5 akad. val., 0,25 ECTS kredito.

VILNIAUS UNIVERSITETAS

PAŽYMĖJIMAS

Nr. MVG-MID-2018-21

VIKTORAS BULAVAS

išklausė Mokslo ir inovacijų departamento
Doktorantūros ir podoktorantūros skyriaus
organizuotus bendrųjų gebėjimų mokymus doktorantams:

- „Mokslinių rezultatų publikavimas“ – 2017 m. lapkričio 16 d., 5 akad. val., 0,25 ECTS kredito;
- „Intelektinės nuosavybės apsauga“ – 2017 m. gruodžio 14 d., 5 akad. val., 0,25 ECTS kredito;
- „Vertės pasiūlymas. Kas tai yra ir kaip jį sukurti?“ – 2018 m. vasario 8 d., 5 akad. val., 0,25 ECTS kredito;
- „Įvadas į *R*“ – 2018 m. kovo 15 d. ir 2018 m. kovo 22 d., 32 akad. val., 1,25 ECTS kredito;
- „Darbas su *LaTeX*“ – 2018 m. balandžio 5 d. ir 2018 m. balandžio 12 d., 32 akad. val., 1,25 ECTS kredito;
- „Mokslo projektai – finansinės priemonės ir paraiškų rengimas“ – 2018 m. balandžio 9 d., 5 akad. val., 0,25 ECTS kredito.

Mokslo ir inovacijų departamento direktorė Vida Lapinskaitė

2018 m. birželio 18 d.

Atlikta: Vasaros mokykla

- 2018 m. liepos 1 – liepos 5 d. papildomai mokytasi tarptautinėje mašininio mokymo vasaros stovykloje, Gdansko technologijų universitete.

Atlikta: Publikacijos

- Bulavas, Viktoras; Dzemyda, Gintautas; Marcinkevičius, Virginijus. An investigation of early cyber threat detection using ensembles of machine learning methods // 9th International workshop on Data Analysis Methods for Software Systems (DAMSS), Druskininkai, Lithuania, November 30 - December 2, 2017. Vilnius : Vilniaus universitetas, 2017. ISBN 9789986680642. p. 9-10. Prieiga per internetą: <https://www.mii.lt/datamss/files/liks_mii_drusk_2017.pdf>.
- Bulavas, Viktoras. „Investigation of network intrusion detection using data visualization methods". The 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University, 2018 (priimtas).

Antrųjų mokslo metų darbo planas

- Uždavinių formulavimas ir metodikos parinkimas 2018 m.
- Teoriniai ir empiriniai tyrimai
- Egzaminai:
  - Atpažinimo teorija – 9 kreditai
  - Optimizavimo teorija, algoritmų sudėtingumas – 7 kreditai
- Dalyvavimas mokslinėje konferencijoje Lietuvoje
- Planuojama parengti vieną mokslinę tyrimų publikaciją konferencijos darbų medžiagoje.

Antrųjų mokslo metų darbo planas

2. Tyrimo metodikos parinkimas:

2.1. Problemų būsimiems eksperimentiniams ir analitiniams tyrimams suformulavimas.

2.2. Uždavinių, skirtų nustatytoms problemoms spręsti, aprašymas.

2.3. Tinkamos tyrimo metodikos parinkimas iškeltiems uždaviniams spręsti.

2.4. Teorinio ir empirinio tyrimų plano parengimas pagal pasirinktą metodiką.

- Planuojama suformuluoti galimas problemas ir hipotezes bei parinkti priemones ir metodus problemų sprendimui (hipotezių patvirtinimui ar atmetimui).

Antrųjų mokslo metų darbo planas

3. Teorinis tyrimas:
3.1. Tinklo įrenginių žurnalinių įrašų apjungimo metodų analizė.
3.2. Metodų, skirtų informacijos saugos anomalijoms duomenyse nustatyti, analizė.
3.3. Tinklo įvykių duomenų tyrimo algoritmų analizė.
4. Empirinis tyrimas:
4.1. Skirtingų algoritmų palyginimas.

- Planuojama parengti vieną mokslinę tyrimų publikaciją konferencijos darbų medžiagoje.

Investigation of network intrusion detection using data visualization methods

THE 59TH INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION TECHNOLOGY AND MANAGEMENT SCIENCE OF RIGA TECHNICAL UNIVERSITY

Intrusion detection scope

- Intrusion detection is a problem within cybersecurity domain.
- Intrusion detection signals malicious activity or policy violations at network or system level.
- Current network intrusion detection appliances utilize three main technics:
  - misuse detection (signature based),
  - anomaly detection,
    - „Anomalies are patterns in data that do not conform to a well defined notion of normal behavior (Chandola, Banerjee, Kumar 2009).“
  - and hybrid.

Intrusion detection problem

- Misuse detection systems use signatures that describe already known attacks and require regular ruleset update.
- Current trend - anomaly detection, based on models, built from normal data, variations from the normal model in the observed data are detected.
- The main advantage with anomaly detection algorithms is that they can detect new forms of attacks, because these new intrusions will probably deviate from the normal behaviour [5].
  - [5] D. E. Denning, "An Intrusion-Detection Model," in 1986 IEEE Symposium on Security and Privacy, 1986, pp. 118–118.
- This way early detection of intrusion becomes possible.

Introduction

- There are numerous sources for network intrusion detection data: for example, network traffic, system host logs, user activity, such as mail or browsing, use of smart devices and similar. All this data comes in big volumes, velocity and variety.
- Analysis of such data is essential for making anomaly detection and intrusion prevention decisions.
- Common data processing steps, following the acquisition of data, are projection, which helps to reduce the number of dimensions, and visualization, which helps observation of distinct features in real time.
- Both steps are required for better understanding of contained intrusion phenomena, such as data theft, malware activity or hacking attempts.

Data Sources

Network data

- The router or switch has the ability to collect IP network traffic as it enters and exits the interface (flows).

Host Data

- The host has the ability to generate system level and user behavior data, usually not obtainable directly from network flows, but related on a temporal axis. Such data would be for example failed login attempts.

All of this data has temporal dimension, which is needed for real life observation of intrusion.

Data sources

- **Primary sources of intrusion detection data, further defined as datasets, are network flows from other network domains and local network, enriched with host-based user behavior and system level content, as shown on Fig. 1 [6], which is needed to detect anomalous behavior and various types of intrusion attacks.**

[6] R. Koch, "Towards Next-Generation Intrusion Detection," in *2011 3rd International Conference on Cyber Conflict*, 2011, February, pp. 151–168.

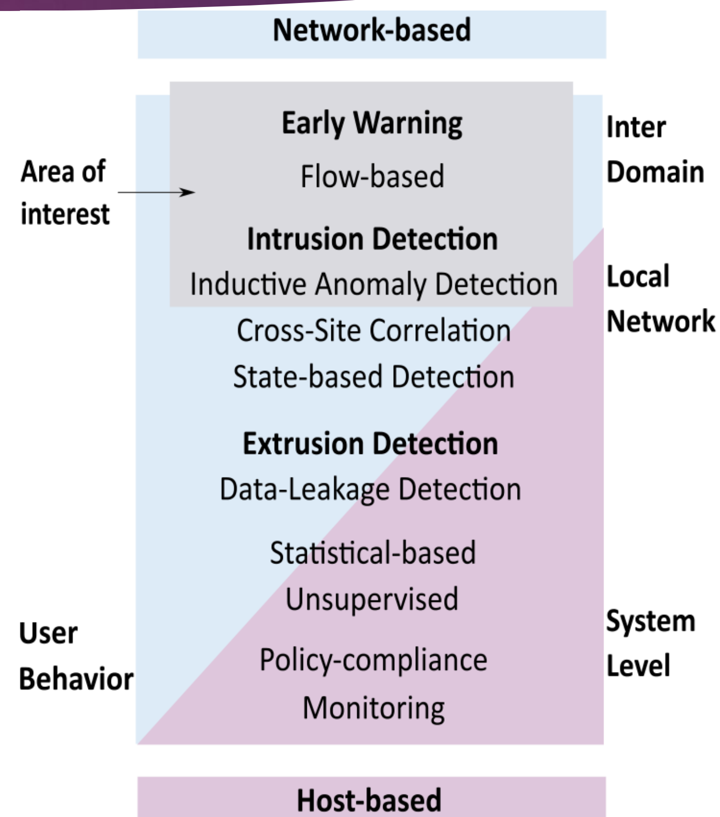


Fig. 1. Layers of intrusion and extrusion detection data [6]

Experiment

- The objective of experiment in this research was to visualise different types of attack data, available in the NSL-KDD dataset.
- Particular attention is drawn to linear projection, in particular principal components analysis, helping to select the most informative dimensions.
- Principal components analysis method, that provides indication of anomalies in network and host data are further reviewed and presented in this paper.
- Decision Tree method is utilized to provide decision criteria for anomaly recognition as intrusion.

Visualization methods

- Dzemyda, Kurasova and Zilinskas [2] classify visualization methods into direct visualization and projection visualization methods:
- 1. Direct visualization methods (when features of a multi-dimensional object are presented in a certain visual form). Using these methods, the selected dimensions of data are presented in a visual form on a two dimensional plane.
  - Direct visualization methods can be further classified as geometric, symbolic and hierarchical.

[2] G. Dzemyda, O. Kurasova, and J. Zilinskas, Multidimensional Data Visualization. Springer, 2012

Principal Components Analysis

- Attention in this analysis is drawn to linear projection, in particular Principal Component Analysis (PCA), as introduced by Hotelling [3], helping to select the most informative dimensions for intrusion detection.
  - [3]H. Hotelling, “Analysis of a complex of statistical variables into principal components,” J. Educ. Psychol., vol. 24, no. 6, pp. 417–441, 1933.
- Brauckhoff, Salamatian and May [29] discuss implementing PCA method for anomaly detection and issue of right number of Principal Components for analysis.
  - [29] D. Brauckhoff, K. Salamatian, and M. May, “Applying PCA for Traffic Anomaly Detection: Problems and Solutions,” in IEEE INFOCOM 2009 - The 28th Conference on Computer Communications, 2009, pp. 2866–2870.
- PCA, together with Decision Tree, can be successfully used for traffic feature extraction and intrusion classification.

Dataset used

- One of most easily accessible for research and education purposes intrusion detection datasets is KDD’99, generated for KDD Cup Contest of 1999.
- It has been updated as NSL–KDD and made available for download at University of Brunswick, Canada.
- The NSL–KDD dataset consists of 41 dimensions [14].

[14] Y. Bouzida and F. Cuppens, “Efficient intrusion detection using principal component analysis,” Proc., 2004.

Features in NSL-KDD

- 1) Basic features: attributes that can be extracted from a TCP/IP connection (ingress interface, source IP address, destination IP address, IP protocol, source port, destination port, and IP type of service).
- 2) Host features: examine only the connections in the past 2 seconds that have the same destination host as the current connection, and calculate statistics related to protocol behaviour, service, etc.

Features in NSL-KDD (continued)

- 3) Service features: examine only the connections in the past 2 seconds that have the same service as the current connection. However, now popular slow probing attacks scan the hosts using a time interval as defined by botnet control centre.
- 4) Content features: unlike most of the DoS and Probing attacks, the R2L and U2R attacks don't have a similar sequential pattern. The R2L and U2R attacks are embedded in the data portions of the packets, and normally represent only a single connection. To detect such attacks, IDS needs specific features in the data portion to recognise as an anomaly, for example a number of failed login attempts. These features are called content features.

Best representing data features for intrusion detection

- Amiri [33], Olusola [34], Zargari [35] and others, based on PCA analysis, proposed methods of selecting the best representing data features of NSL-KDD for intrusion detection:
  - Service, Source bytes, Destination bytes and Destination host error rate.
  - These features explain about 97% of variance. The remaining 37 features explain up to 99,7%, and 80 network features predict 99,97% of attacks.

    - [33] F. Amiri, M. Rezaei Yousefi, C. Lucas, A. Shakery, and N. Yazdani, “Mutual information-based feature selection for intrusion detection systems,” J. Netw. Comput. Appl., vol. 34, no. 4, pp. 1184–1199, 2011.
    - [34] A. A. Olusola, A. S. Oladele, and D. O. Abosede, “Analysis of KDD & apos ; 99 Intrusion Detection Dataset for Selection of Relevance Features Analysis of KDD ’ 99 Intrusion Detection Dataset for Selection of Relevance Features,” vol. I, no. January, pp. 16–23, 2016.
    - [35] S. Zargari and D. Voorhis, “Feature selection in the corrected KDD-dataset,” in Proceedings - 3rd International Conference on Emerging Intelligent Data and Web Technologies, EIDWT 2012, 2012.

Open Source Data Analytics Orange 3

- A simple Orange 3 workflow is used for visualization with ScatterPlot

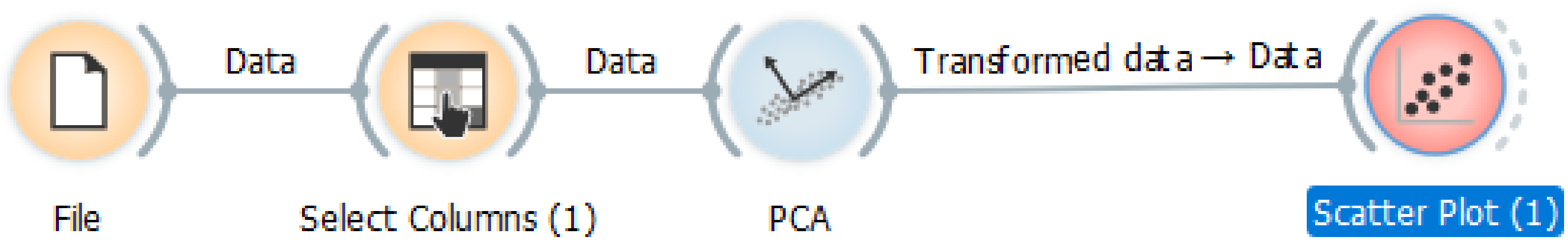


Fig. 4. Principal Component Analysis workflow using Orange 3 software.

Principal Component Analysis of NSL-KDD using Orange 3 software

Fig. 5. Principal Component PC1-PC2

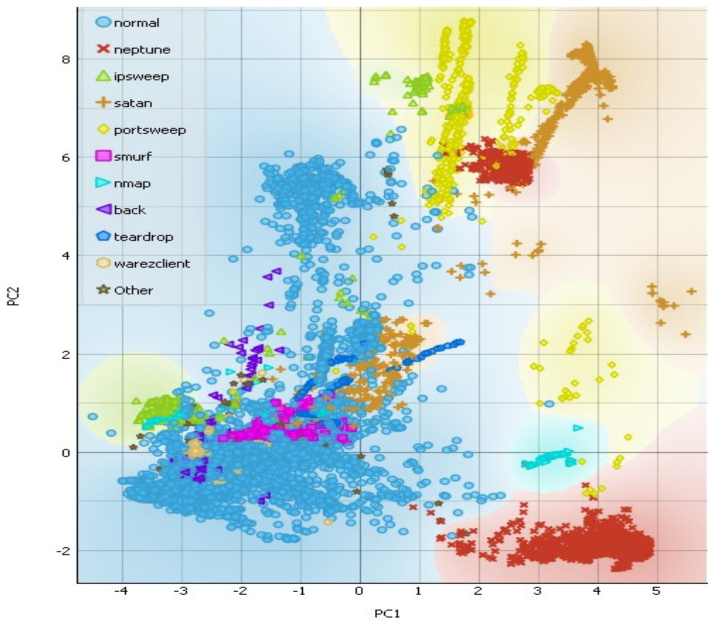
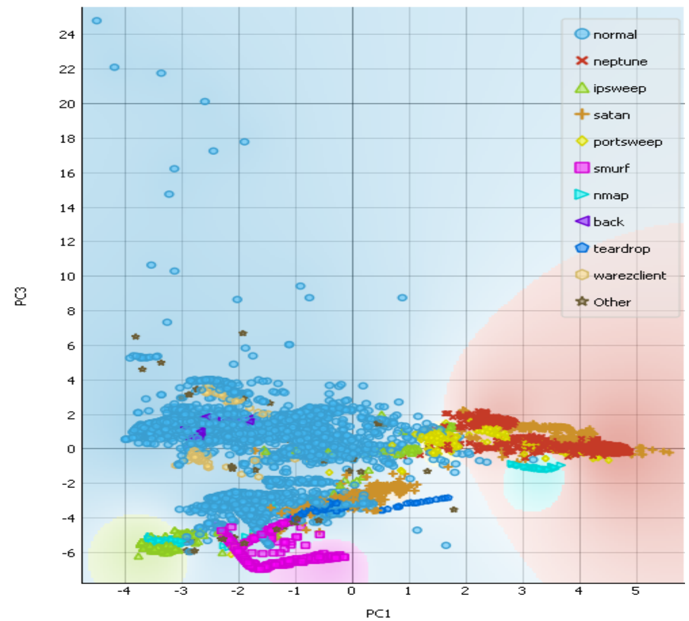


Fig.6. Principal Component PC1-PC3



Open Source Data Analytics Orange 3

- For the purpose of experiment reproducibility, related Orange workflow for PCA analysis with Decision Tree is presented in Fig. 7.

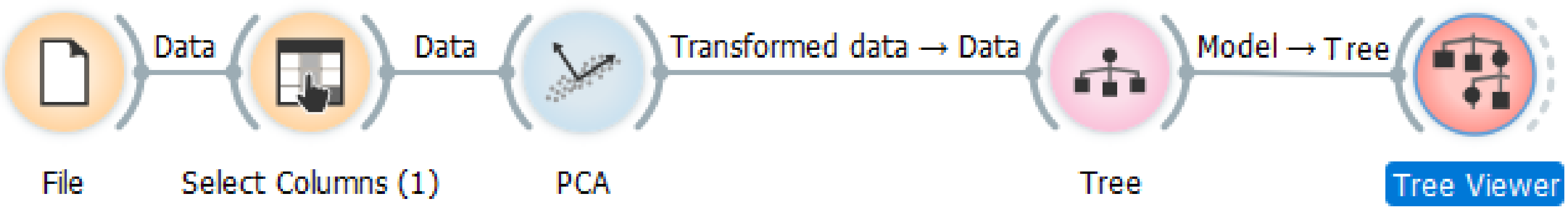


Fig.7. PCA and Decision Tree Analysis workflow using Orange 3 software.

Decision Tree for NSL-KDD

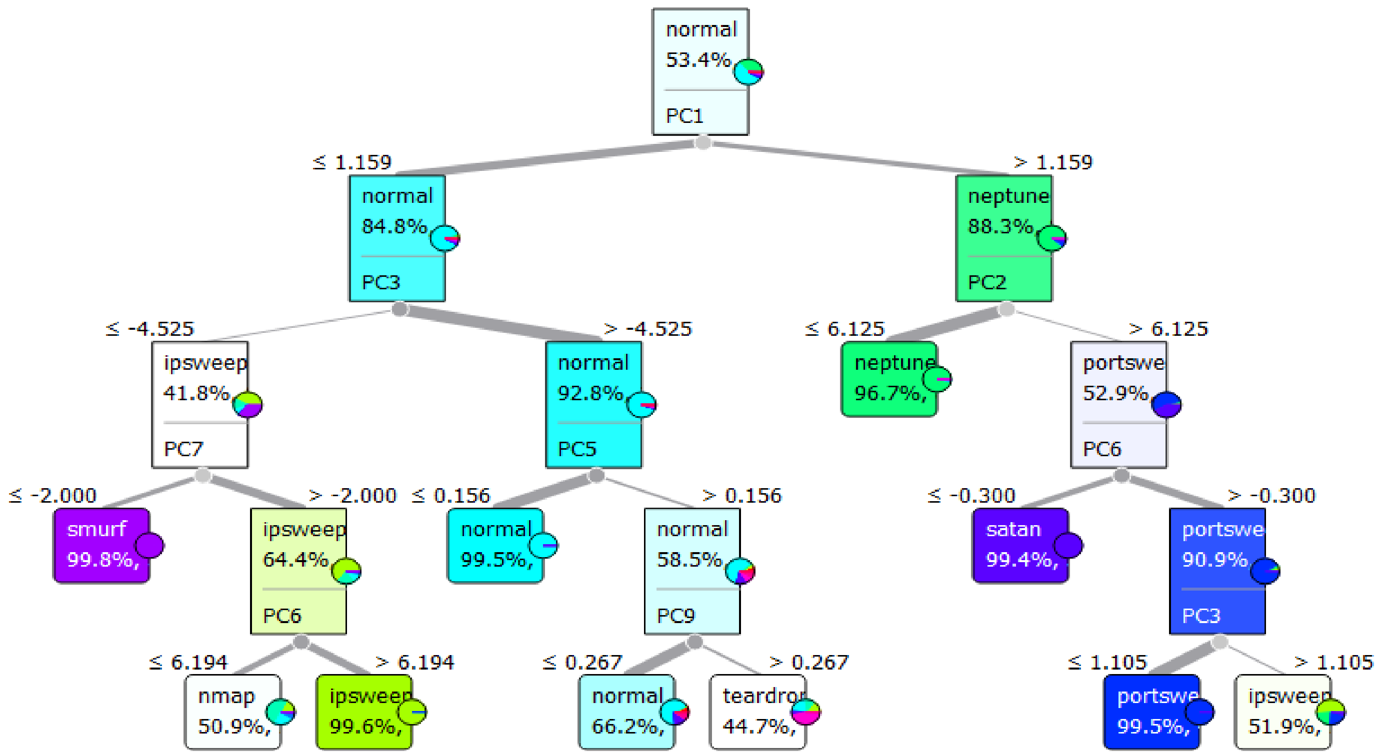


Fig.8. Decision Tree for NSL-KDD data using Orange 3 software.

Findings

- Investigation in this research demonstrates, that combination of PCA and Decision Tree methods allows classification of intrusions such as:
  - smurf,
  - satan,
  - neptune,
  - portsweep,
  - ipsweep
- with probabilities higher than 95% with depth of tree set to 4 and PCA components set to 10.
- Nevertheless, nmap and teardrop intrusions are classified purely, therefore deeper Decision Tree is needed to increase classification accuracy.

Future work

- Future experiment and analysis could be performed using more detailed data source CIC IDS 2017 [13], with ML implemented on open source Tensorflow framework. According to Sharafaldin, Lashkari, and Ghorbani [13], the abovementioned source, enriched with 80 network features, contains 28 informative principal components.
  - [13] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in Proceedings of the 4th International Conference on Information Systems Security and Privacy, 2018, no. January, pp. 108–116.

Future work

- Kim and Reddy [23] demonstrated, that each sample of network flow data could be represented as an image frame or a video stream. For example, the image may represent traffic volume in bytes or packets going to a destination or the traffic between a source and destination pair.
  - [23] S. S. Kim and A. L. N. Reddy, “A study of analyzing network traffic as images in real-time,” in Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies., 2005, vol. 3, pp. 2056–2067.
- Multiple pieces of data can be represented as different colours of an image leading to clear visual presentation and simpler analysis.
- Implement model of conversion of network data into data frame, reproducing algorithms implemented by Kim and Reddy [23].
- Implementing Ensemble and checking if solutions, brought by Hinton et al with Capsule Networks, with learning layers, eliminating need of retraining with feeding of all the data from the beginning.

AČIŪ UŽ DĖMESĮ!

Viktoras Bulavas

E-mail: viktoras.bulavas@itpc.vu.lt