Model of Generative Network Attack

Timur V.Jamgharyan
National Polytechnic University of Armenia
Yerevan, Armenia
t.jamgharyan@yandex.ru

Vahe H. Ispiryan
National Polytechnic University of Armenia
Yerevan, Armenia
vaheispiryan4@gmail.com

Abstract — **This paper considers the application of attacks from the base layer of a data transmission network using a generative model. A making a decision on the expediency of an attack on a network infrastructure based on a generative model has been developed. As a tool for statistical analysis of a generative model was used Pearson's goodness-of-fit criterion. The calculation results demonstrate the principal possibility of applying the Pearson goodness-of-fit criterion for making a decision on the use of a generative model for an attack on a network infrastructure. The traffic exchange point (IXP, Internet Exchange Point) analysis statistics were uses as a dataset to train the generative model.**

Keywords—generative model, generative adversarial networks, complex event, Pearson's goodness-of-fit criteria, network infrastructure, intrusion detection system

I. Introduction

With the development of generative adversarial network (GAN) technology, the issue of using GAN to protect and attack network infrastructure becomes relevant.The typically a network security architecture is built around proven solutions and is governed by security policies [1]-[2]. Depending on the built threat model for each type of network, a risk management model is form and critical objects of the network infrastructure are structured. It should be notes that network security solutions that can successfully applied to one type of network are not always suitable for another type of network [3]. As the boundaries of physical, virtual and software-based networks shrink, the “attack surface” [4] continues to grow. This is especially noticeable in the field of machine-to-machine communications (M2M) and high-load sys-tem [5]. There are different use cases for GANs for different aspects of security, but the problem of evaluating the use of the GAN as a tool for attacks from the network core has not been considered [6]-[7]-[8]. The purpose of the article is to develop a model for assessing the use of generative adversarial network (GAN) for attacks on network infrastructure. To achieve this goal, the following tasks have been solved in the article: the issue of training the GAN for attacking the network infrastructure from the network core level and building a model for making a decision on the use of the GAN

The concepts of building threat models that exist today are mostly deterministic. That is, the scenarios used with certain input data (methods of influencing the network infrastructure), which the attacker performs, generate the corresponding type of attack at the output, which makes it possible to build appropriate templates (rules) for intrusion detection systems.

This property is intrusion detection system a vulnerability of the. Building a secure perimeter against these types of attacks is especially important if an attacker attacks the internal network from the base level (network core) using a generative adversarial network.

It was 2014, Ian Goodfellow introduced the concept [9] of generative adversarial networks, which allow you to create objects similar to the original (images, photographs, music) with varying degrees of veracity. However, in the field of generating false traffic, this approach has not yet become widespread. With the advancement of generative adversarial networking technology, the roles of machine-to-machine attacks and generative adversarial networks (GAN) are becoming increasingly important as attack and defense tools

Accordingly, the problem of assessing and neutralizing the threat proceed from the attacker using the GAN as an attack tool is urgent. This research consider the possibility of using the GAN as attacking tool for the attack from the core level network (attack from the provider side) [10]-[11]-[12].

II. Terms and Definition

A. Basic concepts of generative adversarial networks

Concept of generative adversarial networks invented in 2014 by Ian Goodfellow.

Definition 1: Generative adversarial network (GAN) is an algorithm based on a combination of two neural networks, one of which generates an object, and the other tries to distinguish correct ("real") objects from incorrect ones.

- the generating network *G* (generator) creates (generates) objects of a specified structure,
- the discriminating network *D* (discriminator) draws conclusions about the similarity of the generated and true objects [9]-[13]-[14].

B. Mathematical statistics

Definition 2: Pearson criterion (χ^2) — used to test the hypothesis about the correspondence of the empirical distribution to the assumed theoretical distribution $f(x)$ with a large sample size. The criterion is applicable for any kind of function $f(x)$ [15]-[16].

The criteria is universal.

$$\chi^2 = N \sum_{i=0}^{n} \frac{(P_i^{theor} - P_i^{emp})^2}{P_i^{theor}} \qquad (1)$$

where: $P_i^{theor} = \int_{x_i-1}^{x_i} f(x)dx$ the estimated probability of hitting the *i*-th interval;
$P_i^{emp} = \frac{n_i}{N}$ the empirical value;
n_i — i-th interval elements number sample;
N —summary number of elements.

Definition 3: The zero hypothesis is the default assumption that there is no relationship between two observed events [17].

Definition 4: The competing hypothesis is the opposite hypothesis to the zero hypothesis [17].

C. Information Security

Definition 5: The Dolev-Yao model is a formal model used to prove the properties of interactive cryptographic protocols [18].

Attacker capabilities according to the Dolev-Yao model

- Is an authorized network user;
- Can send messages to any user from any other user;
- Can receive any message transmitted over the network;
- Can become a party receiving messages from any transmitting party.

Restrictions for attacker

- Cannot guess random numbers chosen from a sufficiently large set;
- Cannot decrypt without having a key, or correctly encrypt a message provided that some ideal encryption algorithm is used;
- Cannot find the private key by the public key (when using a public key cryptosystem);
- Access to internal resources such as user memory or hard drive — denied.

Definition 6: A computer attack is an attempt to destroy, disclose, alter, block, steal, unauthorized access to an asset or its unauthorized use [19].

Methods for detecting attacks

- The abuse detection — assumes the presence of attack signatures. The main disadvantage is the inability to detect new or unknown attacks [20].
- The anomaly detection — based on behavior profiles. Any deviation from the profile is an attack. Threat levels are assigned based on the severity of the vulnerability on a scale from 0 to 10 [20]-[21].

D. Big Data

Steps for Data analysis process [22]

- Collecting data
- Preprocessing Data
- Analyzing and Finding Insights
- Insights Interpretations
- Storytelling

III. Attack Parameters

In [23]-[24] various models of development and construction of network infrastructure presented. The research starting point of the model presented in [24]. The attack model is on the assumption that the attacker has the abilities described in the Dolev-Yao threat model. The used by the statistical analysis model (abnormal behavior recording model) will be use as the intrusion detection system. Large deviations of local characteristics from global ones will be a sign of anomalies in the data flow [20]. The possibility of using the generative model is calculate based on the Pearson criterion. For a successful attack, the generative model must process and generate such a volume of data with specified functions so that for the impossibility of detecting an anomaly in the network infrastructure, the χ^2 criterion is less than the specified significance level. In the model do not uses competing hypotheses. If false data entered into the network infrastructure contradicts the null hypothesis, the attack will be detected [16]. An anomaly in the network infrastructure will detected when$\chi^2 \geq \chi^2_0$. It is also possible to use the fact that false network traffic increases the overall entropy as an anomaly detection mechanism. To develop a model of a generative attack on a network infrastructure, the following notation and abstractions, as well as the OSI (Open System Interconnection) model, are used [25]-[26]-[27].

M — true datasets about a system.

$$\left\{m^{(1)}, m^{(2)}, m^{(3)}, \dots\dots, m^{7}\right\} \in M \tag{2}$$

where:

$m^{(1)}$ the physical layer devices dataset,
$m^{(2)}$ the data link layer protocols dataset,
$m^{(3)}$ the network layer protocols dataset,
$m^{(4)}$ the transport layer protocols dataset,
$m^{(5)}$ the session layer protocols dataset,
$m^{(6)}$ the presentation layer protocols dataset,
$m^{(7)}$ the application layer protocols dataset,
Q the datasets, that the GAN must generate.

Condition 1: the false datasets generated by the generative model approximately congruent to the true dataset.

$$Q \cong M \tag{3}$$

Condition 2: for any dataset of protocols, the generative model generates a dataset of approximately congruent protocols with a dataset of that do not contradict Pearson's goodness-of-fit criterion

$$\{M_I^K \cong Q_I^K\} \in X^2 \tag{4}$$

Where: m_i^k — *i*-th sign of *k*-th the protocol of this layer of the OSI model,
q_i^k — *i*-th generated sign of the *k*-th protocol of a given layer of the OSI model.

The result of a generative model attack, model generate a dataset (based on existing data sets), as a result of which the network infrastructure goes into a state of instability [28] (a hidden attack on availability).The generated dataset goes through a set of Bayesian filters and a 5-steps data analysis process [29].

The dataset generated by the generative model must satisfy the following conditions

- will meet the likelihood criterion [28],
- will be not increase the entropy of the system,
- will meet the Pearson's goodness-of-fit criterion,
- will be perceived system as a true data.

IV. Attack Concept

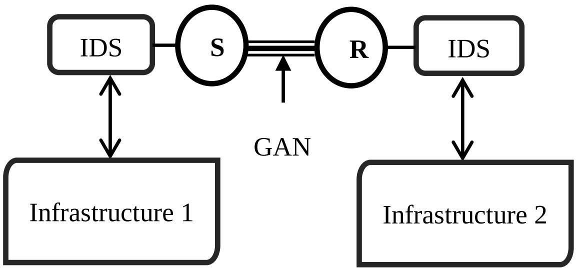

Fig. 1. Scheme of an attack on a network infrastructure using a generative adversarial network

Figure 1 introduces the following notation:

≡ — network core,
IDS — intrusion detected system,
R, *S* — border device (border router, border encrypt device).

A. Attack description

An attacker must carry out man-in-the-middle and traffic analysis attacks and, having captured a large amount of traffic. Based on the generative model, the attacker begins to inject false traffic into the network core. The attack goal is to the availability of S, R devices and train the intrusion detection system to be immune to false traffic for further damage internal network infrastructure unnoticed. This attack on edge devices (S, R) is possible even when using virtual private network (VPN) technology, since the attacker has access to the network core. An attacker based on the generative model can generate a large number of packets that are almost identical to the true traffic, and edge devices will spend computing resources processing headers and only then drop the packet (due to mismatch of secret identifiers: passwords, keys, certificates).

The dataset for input to the system must be formed the basis of the analysis of many factors and their interaction is complex events [30].

B. Decision Algorithm

1. Determination of the set of known true datasets M,
2. Creation of a generative model of datasets, correlating with a given degree of likelihood with M,
3. Making a decision to carry out an attack based on the Pearson's goodness-of-fit criterion.
   - Meets the Pearson criterion — the attack is expedients (input into the system of datasets generated by the generative model);
   - Don`t meet Pearson criterion — the attack is not advisable.

V. Implementation

Assessing the feasibility of carrying out an attack based on a generative model.

Number of empirical distributions — 4.

All calculations carried out taking into account the correction for the continuity of the characteristics to the input/output data. Differences between two distributions can considered significant if χ^2_{emp} reaches or exceeds $\chi^2_{0.05}$ and even more reliable if χ^2_{emp} reaches or exceeds $\chi^2_{0.01}$[29].

Datasets collected through statistical analysis of core network traffic.

Statistical calculations were in the IBM SPSS Statistics software environment [30].

IDS "Snort" and "Suricata" software environments were uses as a "trainable" intrusion detection system [31]-[32].

We introduce the following notation

- N — sampling from datasets,
- $K_{critical}$ — χ^2critical values,
- $\mathrm{f_{emp}}$ — empirical frequencies,
- $\mathrm{f_{theor}}$ — theoretical frequencies,
- $\chi^2 = \frac{(\mathrm{f_{emp}}-\mathrm{f_{theor}})^2}{\mathrm{f_{theor}}}$,
- f—frequency of generation of datasets.

The tables 1, 3 present datasets for creating a generative model. Tables 2, 4 show the results of calculations based on datasets from tables 1, 3 for constructing a generative model.

TABLE 1

First Dataset For Generative Model

$m_1^{(2)}$	$m_1^{(3)}$	$m_1^{(4)}$	$m_1^{(5)}$
13.25	13.30	14.30	12.30
13.40	13.80	14.10	12.60
12.30	12.46	19.80	15.60
12.65	12.50	18.30	15.60
12.60	14.10	12.60	14.30
13.30	14.20	12.30	14.90
14.50	13.20	13.20	21.90

TABLE 2

Results of Calculation for Datasets the First Generative Model

$\mathrm{f_{emp}}$	$\mathrm{f_{theor}}$	$\lvert\Delta \mathrm{f_{emp-theor}}\rvert$	χ^2
13.25	12.31	0.94	0.071
13.30	12.51	0.79	0.05
14.30	13.99	0.31	0.007
12.30	14.34	2.04	0.29
13.40	12.48	0.92	0.068
13.80	12.69	1.11	0.097
14.10	14.19	0.09	0.001
12.60	14.54	1.94	0.259
12.30	13.93	1.63	0.191
12.46	14.16	1.70	0.204
19.80	15.84	3.96	0.99
15.60	16.23	0.63	0.025
12.65	13.67	1.02	0.076
12.50	13.90	1.40	0.141
18.30	15.54	2.76	0.49
15.60	15.93	0.33	0.007
12.60	12.41	0.19	0.003
14.10	12.62	1.48	0.174
12.60	14.11	1.51	0.162
14.30	14.46	0.16	0.002
13.30	12.66	0.64	0.032
14.20	12.88	1.32	0.135
12.30	14.40	2.10	0.306
14.90	14.76	0.14	0.001
14.50	14.54	0.04	0
13.20	14.79	1.59	0.171
13.20	16.53	3.33	0.671
21.90	16.94	4.96	1.452

Result:

$K_{critical1}$ =28,869

$K_{critical2}$ =34,805

$\chi^2_{emp} = 6{,}076$

Since χ^2_{emp} is less than the critical value, the differences between the distributions are not statistically reliable. The use of a generative model using a dataset with the frequencies indicated in table 1 is impractical and the attack will not achieve a result.

That is, that dataset of captured valid (true) traffic is not suitable as input for the generative model.

It is necessary to either increase the sample size or change the base input dataset.

Figure 2 shows a graph of the dependence of the frequencies of empirical and theoretical datasets and the calculation result χ^2

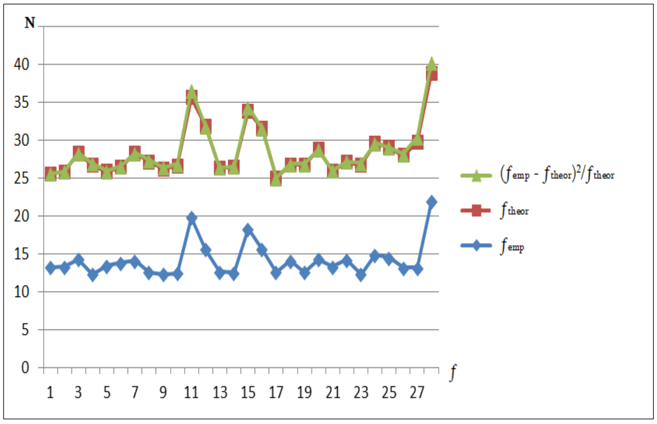

Fig. 2. Plot of the dependence of the frequencies of empirical and theoretical datasets and the result of the calculation χ^2

TABLE 3

SECOND DATASET FOR GENERATIVE MODEL

$m_2^{(2)}$	$m_2^{(3)}$	$m_2^{(4)}$	$m_2^{(5)}$
22.95	32.04	52.60	49.20
6.365	16.125	23.50	48.30
35.30	29.323	46.40	16.40
85.20	34.372	35.60	24.60
7.362	16.656	24.50	13.20
19.23	55.24	49.20	45.50
23.54	22.95	77.30	13.40

TABLE 4

RESULTS OF CALCULATION FOR DATASETS THE SECOND GENERATIVE MODEL

f_{emp}	f_{theor}	$\lvert\Delta f_{emp-theor}\rvert$	χ^2
22.95	33.84	10.89	3.504
32.04	34.99	2.95	0.249
52.60	52.32	0.28	0.002
49.20	35.65	13.55	5.15
6.365	20.35	13.985	9.611
16.125	21.04	4.915	1.148
23.50	31.46	7.96	2.014
48.30	21.44	26.86	33.65
35.30	27.50	7.80	2.212
29.323	28.43	0.893	0.028
46.40	42.52	3.88	0.354
16.40	28.97	12.57	5.454
85.20	38.80	46.40	55.489
34.372	40.11	5.738	0.821
35.60	59.99	24.39	9.916
24.60	40.87	16.27	6.477
7.362	13.32	5.958	2.665
16.656	13.77	2.886	0.605
24.50	20.59	3.91	0.743
13.20	14.03	0.83	0.049
19.23	36.51	17.28	8.179
55.24	37.75	17.49	8.103
49.20	56.45	7.25	0.931
45.50	38.46	7.04	1.289
23.54	29.61	6.07	1.244
22.95	30.61	7.66	1.917
77.3	45.78	31.52	21.702
13.4	31.19	17.79	10.147

Result

$K_{critical1}$ =28,869

$K_{critical2}$ =34,805

$\chi^2_{emp} = 193{,}653$

Figure 3 shows a graph of the dependence of the frequencies of empirical and theoretical datasets and the calculation result χ^2

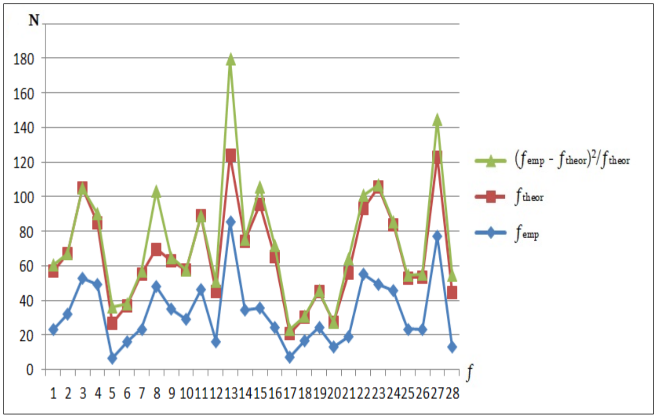

Fig. 3. Plot of the dependence of the frequencies of empirical and theoretical datasets and the result of the calculation χ^2

The since χ^2_{emp} exceeds the critical value, the discrepancies between the distributions are statistically reliable.

That is, the use of a generative model using a dataset with frequencies based on the data indicated in Table 2 is expedient and the attack, other things being equal, will be effective.

VI. Conclusions

The developed model makes it possible to assess the feasibility of using the generative model as a tool for attacking the network infrastructure from the network core level and to make appropriate adjustments to the threat model. These attacks are relevant at the level of peering exchange points between communication providers (ISP, Internet Service Provider). This model also use to "train" intrusion detection and prevention systems in next-generation attacks based on generative models. Based on the model, a developed rules set for open source intrusion detection system Snort and Suricata have been, which, in conjunction with the SIEM (Security Information and Event Management) system, allow you to go into increased security mode when moving from "standard attacks" to attacks using a generative network.

The model allows at the network core level to detect changes in the average entropy value for large accumulated data sets.

The GAN implemented using the Scikit-learn library in the Python programming language.

References

[1] Michael Schwartz, Maciej Machulak. *Securing the Perimeter. Deploying Identity and Access Management with Free Open Source Software*. Apress**. 2018

[2] Jazib Frahim, Qiang Huang. *SSL Remote Access VPNs. An Introduction to designing and configuring SSL virtual private networks.* Cisco Press. Indinapolis, In 46240 USA. Library of Congress Catalog Card Number: 2005923483.

[3] Omar Santos, John Stuppi. *CCNA Securitu. 210-260 Official Cert.Guide.*. Cisco Press. Indinapolis, In 46240 USA. Library of Congress Catalog Card Number:2015938283.

[4] Pratyusa K.Manadhata. *An Attack Surface Metric*. CMU-CS-08-152. School of Computer Science, Carnegie Mellon University. Pittsburg, PA 15213

[5] Micha Gorelick, Ian Ozsvald. *High Performance Python. Practical Performant Programming for Humans*. O'REILLY. Beijing • Boston • Farnham • Sebastopol • Tokyo. Second Edition 2020.

[6] Q. Cheng, S.Zhoy, Y.Shen, D.Kong, C. *WPacket-Level Adversarial Network Traffic Crafting using Sequence Generative Adversarial Networks*21.03.2021.[Online].Available:https://arxiv.org/abs/2103.04794

[7] P.Wang, Z.Wang, F.Ye, X.Chen. ByteSGAN: *A Semi-supervised Generative Adversarial Network for Encrypted Traffic Classification of SDN Edge Gateway in Green Communication Network*. 21.03.2021 [Online]. Available: https://arxiv.org/abs/2103.05250

[8] A.Piplai, S.Sree Laya Chukkapalli, A.Joshi. NAttack! Adversarial Attacks to bypass a GAN based classifier trained to detect Network intrusion.19.02.2020.[Online].Available:https://arxiv.org/abs/2002.08527

[9] Ian J.Goodfellow,J.Pouget-Abadie, M. Mirza, B.Xu, D.Warde-Farley,S.Ozair, A. Courville,Y.Bengio. *Generative Adversarial Networks*. [Online]. Available:https://arxiv.org/abs/1406.2661

[10] Cisco three-tier hierarchical model.
[Online]. Available. http://network.xsp.ru/5_7.php**.**.

[11] Sudhasan Ravichandiran.*Deep Reinforcement Learning with Python. Master classic RL, deep RL,distributional RL, inverse RL, and more with open AI Gym and TensorFlow*.Second Edition.. Packt. Birmingham-Mumbai. September 2020

[12] Suresh Kumar Mukhiya, Usman Ahmed.*Hands-On. Exploratory Data Analysis with Python*.Perform EDA techniques to understand, summarize, and investigate your data. Packt. Birmingham-Mumbai. March 2020

[13] Ahmed Fawzy Gad. Fatima Ezzahra Jarmouni. *Learning and Neural Networks with Python*™. A Practical Guide
. ACADEMIC PRESS.Elsevier

[14] Ronald T. Kneusel. *Practical Deep Learning. A python – based introduction*. San Francisco. 2021

[15] Pearson goodness-of-fit criteria
.[Online].Available:https://help.fsight.ru/ru/mergedProjects/lib/05_statisics/modelling_chitest.html.

[16] Joel Grus. *Data Science from Scratch*. O'REILLY. "BXV-Peterburg" 2021

[17] Synergu university electronic library. Statistika.
.[Online].Available.
http://www.e-biblio.ru/book/bib/10_statitika/tv_i_ms/book/docs

[18] Wenbo Mao.*Modern Cryptography: Theory and Practice.*
Prentice Hall.

[19] The FSTEC Russia official website. [Online]. Available https://bdu.fstec.ru/ubi/terms/terms/view/id/2.

[20] O.I. Shelukhin, D. Zh. Sakalema, A.S. Filinov. *Detection of intrusions into computer networks.*

[21] The FSTEC Russia official website. [Online]. Available https://bdu.fstec.ru/ubi/terms/terms/view/id/56

[22] Avinash Navlani. Armando Fandango. Ivan Idris. Packt. Birmingham-Mumbai. *Python Data Analysis. Third Edition. Perform data collection, data processing, wrangling, visualization, and model building using Python*.. February 2021

[23] Andrew S. Tanenbaum, David Wetherall. *Computer Networks.. 5th Edition.* PEARSON Prentice Hall. Piter .2012

[24] Cisco network design. SEDP- Network Foundation Design.
[Online].Available:.https://www.cisco.com/c/dam/en/us
/td/docs/solutions/Enterprise/Small_Enterprise_Design_Profile/chap2sba.pdf

[25] ISO/IEC 7498-l Second edition 1994-l l-l 5 Corrected and reprinted 1996

[26] *Host extensions for IP Multicasting*. RFC 1112. August 1989

[27] *Essnetial Tools for the OSI Internet*. RFC 1574.february 1994

[28] Generative Deep Learning. Teaching mashine to paint, write, compose, and play. David Foster. Beijing-Boston- Farnham-Sebastopol-Tokyo. O`Reilly 2019.

[29] Igor Kotenko, Igor Saenko, D.Kotenko.*Methods and tools for attack modeling in large computer networks: state of the problem*. Artisle in SPIRAS proceedings. March 2014.

[30] M.A. Pleskunova, L.V. Korchemkin. *Probability theory*. Handbook.. Yekaterinburg. Publishing house of the Ural University. 2017 November

[31] Service for automated calculation and description of statistics.
[Online].Available https://stanly.statpsy.ru/all/acc
ounts/login/next=/all/correlation_pearson/steps/1/

[32] IBM corporated official website software section:
[Online].Available.https://www.ibm.com/analytics/spss-statistics-software

[33] Snort IDS/IPS official website.
[Online].Available https://www.snort.org/

[34] Suricata IDS/IPS official website.
[Online].Available https://suricata.io/