**МИНОБРНАУКИ РОССИИ**

Федеральный исследовательский центр «Информатика и управление» РАН
Национальный комитет при президиуме РАН по распознаванию образов и анализу изображений
Институт информационных технологий Национальной академии наук Азербайджана
Институт проблем передачи информации им. А.А. Харкевича РАН
Издательство «Наука и технологии»
Национальный исследовательский Томский государственный университет
Федеральный исследовательский центр «Карельский научный центр РАН»
Федеральное государственное бюджетное образовательное учреждение высшего образования «Юго-Западный государственный университет»

# ОПТИКО-ЭЛЕКТРОННЫЕ ПРИБОРЫ И УСТРОЙСТВА В СИСТЕМАХ РАСПОЗНАВАНИЯ ОБРАЗОВ И ОБРАБОТКИ ИЗОБРАЖЕНИЙ

**Распознавание – 2021**

Сборник материалов XVI Международной научно-технической конференции

14–17 сентября 2021 года

Редакционная коллегия:
С. Г. Емельянов, В. С. Титов (отв. ред.),
Т. А. Ширабакина, Э. И. Ватутин,
В. С. Панищев

Курск 2021

УДК 621.383.68.3: 681.785
ББК В 338.4
О 66



О 66 **Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений. Распознавание – 2021**: сб. материалов XVI Междунар. науч.-техн. конф. / ред. кол.: С. Г. Емельянов, В. С. Титов (отв. ред.) [и др.]; Юго-Зап. гос. ун-т. – Курск, 2021. – 284 с.

ISBN 978-5-7681-1520-3

Сборник содержит материалы XVI Международной научно-технической конференции «Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений» (Курск, 14–17 сентября 2021 г.), целью которой является ознакомление с имеющимися достижениями по созданию оптико-электронных приборов, систем и внедрение информационных технологий в научные исследования, учебный процесс и промышленность, а также координация по эффективному их применению в системах распознавания образов и обработки изображений.

Сборник предназначен для научных сотрудников, преподавателей, аспирантов и студентов вузов.

*Издание осуществлено с авторских оригиналов. Редакция не несет ответственности за ошибки авторов.*

Материалы для публикации одобрены программным комитетом Международной научно-технической конференции.

УДК 621.383.68.3: 681.785
ББК В 338.4

ISBN 978-5-7681-1520-3



## СОДЕРЖАНИЕ

At the moment demographers leverage online data to study the three main components of demographic change: fertility, mortality and migration.

But demographers that used digital data face some problems related to access, representativity, and ethics.

Researchers often don't access for data since internet companies, unlike governments, are not obliged to share data from their platforms. Furthermore, digital sources are rarely representative of larger populations in the way that randomized surveys are. Selection bias poses even more intractable problem for the study of demography with digital data. At the core of the problem is the "digital divide", a fundamental inequality of the 21st century, with individuals across the world having widely different levels of access to the Internet.

Ethical issues must be a primary concern when designing demographic studies using digital data. Social scientists need to adhere to ethical research practices, particularly as the privacy of users is constantly threatened in the online world.

Digital data will fully realize its promise only if researchers are should have the basic skills of data harvesting, processing, and analysis.

## REFERENCES

1. Weber I., State B. Digital Demography // Processing of the International World Wide Web Conference Committee (IW3C2). Perth, Australia, 2017. April 3–7. P. 935–939.

2. Demography in the Digital Era: New Data Sources for Population Research / D. A-Gutierrez, S.Aref [et al.] // Arbia G., Peluso S., Pini A., Rivellini G. (eds.). Book of short Papers SIS2019. Pearson, 2019.

3. Billari F., Zagheni E. Big data and population processes: a revolution? // Petrucci A., Verde R. (eds.) Statistics and Data Science: new challenges, new generations. Proceedings of the Conference of the Italian Statistical Society. Firenze University Press. Florence (Italy), 2017. 28–30 June. P. 167–178.

UDK 004.056.5

***Y. N. Imamverdiyev[1], F. J. Abdullayeva[1]***
e-mail: 1yadigarimam@gmail.com, 2a_farqana@mail.ru
[1] *Institute of Information Technology of ANAS, Baku, Azerbaijan*

## CONVOLUTIONAL NEURAL NETWORK FOR DETECTING APPLICATION LAYER DISTRIBUTED DENIAL OF SERVICE ATTACKS

*Distributed Denial of Service (DDoS) is one of the main threats to information security. Application layer DDoS attacks (AL-DDoS) can be organized against many different applications. Many of these attacks target HTTP, in which case their goal is to consume the resources of web services. This study proposes an approach to detecting AL-DDoS attacks.*



The architecture of an automated diagnostic system for detecting AL-DDoS attacks is illustrated in figure.

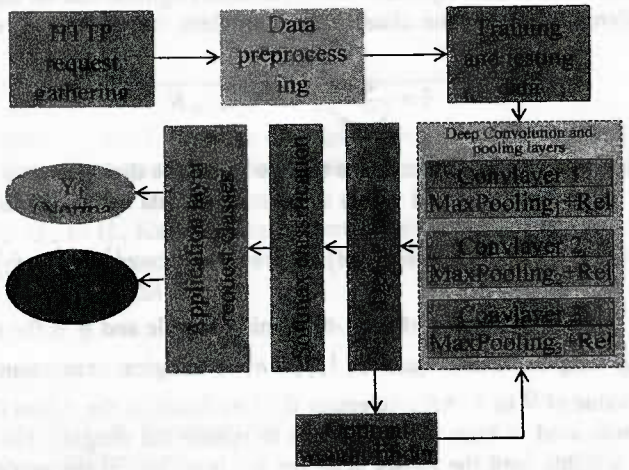


Fig. Automatic diagnostic system for detecting AL-DDoS attacks

In the proposed approach, the CNN model consists of a convolution layer, a Max-pooling layer, a flatten layer, and the last fully connected layer. In general, CNN's input consists of raw data in the form of $I \in R^k$, and its output consists of the results of the classification $\hat{y}$.

$$\hat{y} = act(FCN(Flatt(pool(\mathrm{Re}\,LU(conv(I)))))).$$

The conv layer of the CNN model uses a number of filters to convolve with the input data. The $i$-th vector of the I vector is calculated as

$$I_i = \left[I_{1+(i-1)d}, I_{2+(i-1)d}, \ldots, I_{m+(i-1)d}\right]^T \quad \left(i = 1, 2, \ldots, \frac{k-m}{d}+1\right).$$

The function of 1D convolution is to calculate the product between the vector H and the data vector $I_i$ obtained from the raw data: $S_i = I_i \times H + b = \sum_{j=1}^{m} I_{j+(i-1)d} H_j + b$, where, $H_j$ is the $j$-th element of the vector $H$, $j = 1, 2, \ldots, m$. Relu (Rectified Linear Unit) was used as an activation function in

each convolution layer of the model: $U_i = \mathrm{Re}lu(S_i) \triangleq \max(0, S_i)$. Maxpooling was used in the model as a combination operation: $pool(U_i) := \max_{l=1}^{p} U_{l+(i-1)e}, \quad \forall i = 1, 2, \ldots, \frac{k-m}{d} + 1$, where $p$ is the size of the pulling, $e$ is the stride size. Since the classification problem is solved in the article, softmax was used as an output activation function (sigmoid can be used in regression problem). Thus, for the classification problem, the softmax is expressed as follows:

$$\hat{y} = \frac{e^{O_n}}{\sum_{j=1}^{N} e^{O}}, \quad n = 1, 2, \ldots, N.$$

The cross-entropy function was used to minimize the difference between the predicted values and the actual values in the training data and is defined as follows:

$$L_{\text{crossentropy}} = -\frac{1}{q}\sum_{i=1}^{q}\sum_{n=1}^{N} 1\{y_i = n\} \log \hat{y}_i + \left(1 - 1\{y_i = n\}\right) \log\left(1 - \hat{y}_i\right),$$

where $y_i$ is the output value of the $i$-th training sample and $q$ is the total number of training samples. In the equation, $1\{y_i = n\}$ is a logical expression that always returns a value of $0$ or $1$. After selecting the loss function, the Adam Optimization function was used to train the parameters to update the weights. Here, CNN updates the weights until the model achieves the least loss at the predefined maximum iteration. The CSIC2010 and CSE-CIC-IDS2018 datasets were used for the experiments. Here, the effectiveness of the method was evaluated on the basis of Accuracy, Precision, recall and f1-score metrics, and the results are included in table.

Evaluation CNN efficiency on CSE-CIC-IDS2018 DDoS and CSIC 2010 datasets

| | Class | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| CSE-CIC-IDS2018 DDoS dataset | Benign (0) | 0.9853 | 0.9923 | 0.9921 | 0.9921 |
| | DDoS (1) | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| | DoS (2) | 0.9931 | 0.9903 | 0.9901 | 0.9944 |
| CSIC 2010 dataset | Anomalous Traffic (0) | 0.6331 | 0.9201 | 0.6311 | 0.7531 |
| | Normal Traffic (1) | 0.9911 | 0.9121 | 0.9924 | 0.9521 |

As can be seen from Table 1, the model was able to detect Anomalous Traffic (0) class data with low efficiency when tested on CSIC 2010 dataset data. Thus, the model over the accuracy, precision, recall, and F1-score metrics achieved 0.63, 0.92, 0.63, 0.75 values respectively. In the recognition of Normal Traffic (1) data, the model showed good results and achieved 0.99, 0.91, 0.99, 0.95 for these metrics, respectively. When the model was tested on the CSE-CIC-IDS2018 DDoS dataset, high results were obtained for almost all metrics.



*This work was supported by the Science Development Foundation under the President of the Republic of Azerbaijan – Grant No. EIF-BGM-4-RFTF-1/2017-21/08/1.*

1. Singh K. J., De T. Analysis of Application Layer DDoS Attack Detection Parameters Using Statistical Classifiers // Internetworking Indonesia. 2017. N 9(2). P. 23–31.

2. Liao Q., Li H., Kang S., Liu C. Feature extraction and construction of application layer DDoS attack based on user behavior / 33rd Chinese Control Conference (CCC), 2014. P. 5492–5497.

UDC 004.052.42

***T. H. Kazimov*[1], *T. A. Bayramova*[1]**
e-mail: toma_b66@mail.ru
[1] *Institute of Information Technology of ANAS, Baku, Azerbaijan*

## ABOUT A METHOD FOR EVALUATING THE DEGREE OF SOFTWARE COMPLEXITY INTRODUCTION

*Application of program code metrics allows professionals who work on the project to evaluate various feauters of existing or to be created software, to predict the scope of work, quantitatively characterize these or other project solutions, to evaluate quality of prepared systems, complexity and realiability of software.*

**Problem setting.** Different metrics (especially quantities) that are used when evaluating the efforts and efforts of the program employee are of a recommendation nature. Because some workers deliberately reduce or inflate these indicators. Therefore, an assessment of the complexity of the task facing an employee can play a significant role in solving this problem. For example, it may take days, weeks, and sometimes months, to find a bug in any complex program, but its correction results in a change of one line of program code [1, 2].