Anomalies detection in social services data in the sphere of digital economy

P V Khripunov\textsuperscript{1}, E Y Minaev\textsuperscript{1}, V I Protseiko\textsuperscript{1,2}, N S Davydov\textsuperscript{1,2}, A V Nikonorov\textsuperscript{1,2}

\textsuperscript{1}Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086
\textsuperscript{2}Image Processing Systems Institute of RAS - Branch of the FSRC ”Crystallography and Photonics” RAS, Molodogvardiejskaya street 151, Samara, Russia, 443001

e-mail: eminaev@gmail.com, artniko@gmail.com

Abstract. This article addresses the study of the anomaly and fraud detection problem in the data from social services. The problem of detecting anomalies is extremely relevant for data-driven processes in the digital economy. In this paper, we propose a two-step approach for the detection of anomalies using auto-encoders and the conjugacy indicator. An experimental study of the efficiency of the proposed algorithms was conducted using open-access data set.

1. Introduction
Due to a rapid advancement of digitalization and data-driven approach in the social and economic sphere, public administration and business, the importance of fraud and data anomalies detection task for huge information systems is dramatically increased [1], [15]. On the one hand, the digitalization of public services and business processes can significantly accelerate and simplify access to the necessary information, on the other hand, with an increase number of information systems and links between them, control over the legality of access to data is often weakened, which can lead to undesirable distortions. Moreover, an increasing number of users in information systems lead to an increasing possibility of errors and inaccuracies at the stage of data collection [2].

2. Anomaly analysis in social and government services in the digital economy
Data mining methods for anomaly detection in information services are divided into six main categories: classification methods, clustering methods, regression, outlier detection, visualization and forecasting [3]. Each of these categories includes specific methods. For example, neural networks and support vector machines are used to classify data, the K-means method is used to cluster data. In addition, data mining includes many methods from other areas, such as statistics, machine learning, pattern recognition, databases and data warehouses, information retrieval, visualization, high-performance computing and other application areas [4]. Recently, fraud detection has combined an anomaly detection approach with data mining approach [5].

The method for detecting anomalies or outliers relies on behavioral profiling methods that model the behavior of each person and track any deviations from the norm. Anomaly-based fraud detection
3. Anomaly detection based on autoencoder networks

Autoencoders are neural networks whose purpose is to learn the identity mapping subject to a number of constraints imposed on its architecture. One of these constraints may be fewer internal neurons than external ones. Activations obtained on the smallest layer allow us to represent the initial data in a compressed form and are widely used in applications for further machine processing. Such activations are called encoder output values. Autoencoders having equal or more neurons in the inner layers are also of interest, with a certain regularization or the presence of a penalty in the minimization function. Contractive autoencoders allow encoders to be less sensitive to weak changes in training sample data due to regularization consistent with the Frobenius norm of the Jacobi matrix of encoder activations. By limiting the number of non-zero activations of the auto-encoder per input sample, the learning process allows you to get an encoder that returns sparse vectors of activation values. Adding Gaussian noise to the input vectors, or to the activation of the inner layers, leads to the training of weights with smoother gradients. Such noise-resistant auto-encoders are called denoising.

Based on the autoencoder, a classifier that determines whether it belongs to the class on which it was trained can be built. The rule of such classifier depends on the selected threshold value applied to the error between the input and reconstructed data. Changing the threshold value allows you to get the ratio of the number of false-positive and false-negative responses that arise when it is impossible to strictly separate classes. An autoencoder can be trained to obtain an encoder that maps the original data space to a space in which the quality of classical classification methods is improved, for example, to the space where the data becomes linearly separable. The architecture and analysis of the experimental results of such a two-stage classifier will be presented later in the work.

4. Accuracy improvement using conjugacy indicator

As the second stage of classification, it is proposed to use the conjugacy indicator. To construct the classifier, we will use the approach described in [7], [10]. For each k-class from M training vectors $x_{j}(k)$, $j = 1, M$, $k = 1, K$ which is belonging to him, the matrix with dimensions $N \times M$ is creating:
\[ X_k = \left[ x_1 (k), x_2 (k), \ldots, x_j (k), \ldots, x_M (k) \right], k = 1, K \]  

(1)

And matrix \( N \times N \) for each \( k \)-class is calculated:

\[ Q_k = X_k \left[ x_k^T x_k \right]^{-1} x_k^T, \quad k = 1, K, \]  

(2)

which we called resolving matrix.

At the recognition stage, the decision on the ownership of the vector \( x_j \) to \( m \)-class is accepted, when

\[ R_m (x_j) = \max_{k = 1, K} R_k (x_j), \]

(3)

where \( R_k (x_j) = x_j^T Q_k x_j \left( x_j^T x_j \right)^{-1} \). \( k = 1, K \) — conjugacy indicators of the current vector \( x_j \) with each recognized class.

In this method, information about classes is contained in matrices \( Q_k \mid k = 1, K \), computed using matrices \( X_j (M) \). In the paper [8] to form these matrices, it was proposed to use a small number of training vectors that form the support subspaces of classes. Due to this property, the conjugacy indicator can be used in the case of a small training sample, which often accompanies the problem of classifying anomalies.

5. Results and experiments

In this work, the outputs of the inner layer of the autoencoder are used as the initial vectors for which the classifier is trained. In this work, we used an autoencoder with 4 layers, the two internal layers of which were twice as large in the number of neurons — 58 than the two external ones — 29. During the analysis of the data set, a small number of fraudulent transactions were found in the data, which were very closely located next to regular transactions. The increase in dimension at the first stage of the two-stage classification made it easier to separate the two classes. As functions of layer activation, \( \tanh \) and LeakyRELU functions alternating from the first to the last layer were used.

The following parameters were chosen for training: optimization method — RMSProp [9], the number of stages is — 50, the sample size of the iteration of the optimization method is 32, the loss function is the mean square deviation. Training of the autoencoder was carried out on ordinary transactions, accounting for 80\% of the total number of transactions. The test data contained all fraudulent transactions and about 20\% of ordinary transactions.

Before training, we performed data preprocessing. The column with the transaction time was excluded, since its presence did not improve the classification quality and worsened the separability of fraudulent transaction groups in the two-dimensional t-sne projection. A column was also normalized with the Amount translation amount for a comparable range of values with other attributes.

The following classification metrics were used: AUC and PR-AUC. Due to the imbalance of the data set and priority in identifying a poorly presented class of fraudulent transactions, PR-AUC will be used as the main metric. High AUC values will show how well the two classes separate.

5.1. Experimental dataset description

Dataset [10] contains the sequence of transactions of European cardholders for two days in September 2013. Transactions are divided into two classes: 492 fraudulent and 284,315 legal. The data set is highly unbalanced — fraudulent transactions account for only 0.172\% of the total. For reasons of personal data privacy, the table contains only numerical values that are the result of the transformation by the method of the main components of the original transaction logs. The exception is two columns: the amount of transfer and the time in seconds, counted from the beginning of the first transaction from the data set.
5.2. Results
The use of non-linear activation of LeakyReLU with a slope coefficient of 0.3 made it possible to increase the PR-AUC in comparison with the classical ReLU from 0.47 to 0.72.

Figures 1a and 1b show the results of a classifier built exclusively on a trained auto-encoder. The classification metrics for AUC and PR-AUC were 0.96 and 0.72, respectively. Figure 2 shows the visualization of correctly and incorrectly classified cases of falsification in the space of pre-processed features. You can see that the classification errors of the autoencoder make up a fairly dense cluster. An additional analysis showed that the missing falsifications have a strong intersection of the distribution of characteristic values with the distribution of the corresponding characteristics of ordinary transactions. Figure 3 shows the error matrix.

![Figure 1](image1.png)

*Figure 1. The results of the one-stage classifier built on the basis of the auto-encoder.*

![Figure 2](image2.png)

*Figure 2. Correctly and erroneously classified cases of falsification in the space of pre-processed features.*

![Figure 3](image3.png)

*Figure 3. Matrix of classification errors by one-stage classifier. 0 corresponds to the class of ordinary, and 1 to the class of fraudulent transactions.*
To improve the recognition of 83 anomalies that form a cluster on Figure 2, which could not be classified by an autoencoder and reduce false 534 responses, the second recognition stage was applied using the criterion for conjugate indicator. As the initial vectors by which the classifier was trained, intermediate outputs of the inner layer of the autoencoder of 58 elements were used. Due to the construction of support subspaces by 10% of incorrectly recognized vectors, as a result of the second stage of the classification, it was possible to reduce the number of unrecognized fraudulent transactions to 25 and the number of false positives to 157 transactions. Thus, the proposed approach, which consists in refinement of the detector based on autoencoders through the application of the contingency criterion, has proved its effectiveness on a test data set.

6. Conclusion
In this work, the problem of detecting fraud in social services data was studied. Despite the relevance of creating effective algorithms for fraud detection in data generated by the digital economy, there is an extremely small number of data sets in public domain that allow such algorithms to be validated. Studies have shown that the classic approach based on autoencoders allows for the detection of fraud, but the accuracy of such detection is low. An increase in the accuracy was made possible by our additional classification stage based on the contingency criterion. Testing the effectiveness of the proposed approach to detecting anomalies on other data sets obtained in social systems, as well as validating the algorithm on synthesized data with previously known characteristics, is the subject of further research.

7. References
[1] Dong W, Liao S and Zhang Z 2018 Leveraging financial social media data for corporate fraud detection Journal of Management Information Systems 35(2) 461-487
[2] Phua C, Lee V, Smith K and Gayler R 2010 A comprehensive survey of data mining-based fraud detection research Preprint arXiv:1009.6119
[3] Abdallah A, Maarof M A and Zainal A 2016 Fraud detection system: A survey Journal of Network and Computer Applications 68 90-113 DOI:10.1016/j.jnca.2016.04.007
[4] Han J, Pei J and Kamber M 2011 Data mining: concepts and techniques (Elsevier)
[5] Sasirekha M 2012 A defense mechanism for credit card fraud detection Int. J. Cryptogr. Inf. Secur. 2(3) 89-100
[6] Akhilomen J 2013 Data mining application for cyber credit-card fraud detection system Industrial Conference on Data Mining 218-228
[7] Zherdev D A, Kazanskiy N L and Fursov V A 2015 Object recognition in radar images using conjugation indices and support subspaces Computer Optics 39(2) 255-264 DOI: 10.18287/0134-2452-2015-39-2-255-264
[8] Zherdev D A, Kazanskiy N L and Fursov V A 2014 Object recognition by the radar signatures of electromagnetic field scattering on base of support subspaces method Computer Optics 38(3) 503-510
[9] Hinton G 2014 RMSProp URL: https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slideslec6.pdf (accessed: 2019-09-06)
[10] Credit Card Fraud Detection Dataset URL: https://www.kaggle.com/mlg-ulb/creditcardfraud (accessed: 2019-09-06)
[11] Bibikov S A, Kazanskiy N L and Fursov V A 2018 Vegetation type recognition in hyperspectral images using a conjugacy indicator Computer Optics 42(5) 846-854 DOI: 10.18287/2412-6179-2018-42-5-846-854
[12] Nikonorov A, Bibikov S, Yakimov P and Fursov V 2014 Spectrum shape elements model to correct color and hyperspectral images 8th IAPR Workshop on Pattern Recognition in Remote Sensing
[13] Nikonorov A, Bibikov S, Myasnikov V, Yuzifovich Yu and Fursov V 2016 Correcting color and hyperspectral images with identification of distortion mode Pattern Recognition Letters 83(2) 178-187
[14] Bibikov S, Zakharov R, Nikonorov A, Fursov A and Yakimov P 2011 Detection and color correction of artifacts in digital images Optoelectronics, Instrumentation and Data Processing 47(2) 226-232

[15] Abdallah A, Maarof M and Zainal A 2016 Fraud detection system: A survey Journal of Network and Computer Applications 68 90-113

Acknowledgements
Algorithms were designed and developed with the support of RFBR grants (projects 19-29-01235-мк, 16-29-11744-оф_м, № 16-29-09528- оф_м, № 17-29-03112- оф_м, № 18-07-01390-А), experimental studies - in the framework of the state assignment of the IPSI RAS - a branch of the Federal Scientific-Research Center "Crystallography and Photonics" of the RAS (agreement № 007-Г3/Ч3363/26).