ISSN 1815-4840
Himičeskaâ tehnologiâ. Kontrol′ i upravlenie
CHEMICAL TECHNOLOGY. CONTROL AND MANAGEMENT
2020, №3 (93) pp.59-65
International scientific and technical journal
journal homepage: https://uzjournals.edu.uz/ijctcm/
Since 2005

UDC 004.056

ELECTRONIC SPAM FILTERING BASED ON NEURAL NETWORKS

Bekmuratov Tulkun Fayzievich[1], Botirov Fayzullajon Bakhtiyorovich[2], Haydarov Elshod Dilshod ugli[3]

[1]SIC ICT of TUIT, Tashkent, Uzbekistan,
E-mail: bek.tulkun@yandex.ru, Phone: +998 71 234 07 92;
[2]TUIT named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan,
E-mail: botirov_fz@mail.ru, Phone: +998 97 751 16 9;
[3]TUIT named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan,
E-mail: elshodhaydarov1881@gmail.com, Phone: +998 97 306 99 95

Abstract: *This article analyzes the problem of filtering spam messages and addressing spam messages, Bayesian theorems based on artificial intelligence, LVQ algorithms (LVQ learning vector quantization) and a filtering scheme for systems based on neural networks. The direct construction of an effective neural network model of spam filtering using database recognition technology is considered. The parameters of access to the neural network how to include predefined statistical and non-statistical attributes of messages are given. The structure of neural network technology for classifying emails is also considered. The procedure for analyzing incoming data using the tool included in the analytical platform Deductor Studio 5.3 is described. as a result, a training kit is obtained that is suitable for use.*
Keywords: *spam, artificial intelligence, SOM, LVQ algorithm, information security.*

Аннотация: *Спам хабарни фильтрлаш масаласи ва бу масалани ҳал қилиш учун спам хабарларнинг сигнатуралари, сунъий интелект асосида қурилган Байес теоремалари ва LVQ (LVQ- learning vector quantization) алгоритмлари таҳлил қилиниб, нейрон тармоқ асосида спам хабарларни фильтрлаш тизимининг схемаси кўриб чиқилган. Маълумотлар базасини, таниб олиш технологиясидан фойдаланиб, амалга ошириладиган спамни фильтрлашни самарали нейрон тармоқли тўгридан-тўгри қурилиши кўриб чиқилган. Хабарларнинг маълум статистик ва ностатистик атрибутларини киритиш каби нейрон тармоқларга рухсат параметрлари келтирилган. Яна элекрон хатларни таснифлашнинг нейрон тармоқли технологиялари тузилмаси кўриб чиқилган. Фойдаланиш учун яроқли бўлган натижавий синов комплект олинадиган Дедустор Студио 5.3 аналитик платформасига кирувчи иструментдан фойдаланиб, кирувчи маълумотларни таҳлиллаш амали тавсифланган.*
Таянч сўзлар: *спам, сунъий интелект, SOM, LVQ алгоритми, ахборот хавфсизлиги.*

Аннотация: *Рассмотрены задачи фильтрации спам-сообщений и адресации спам-сообщений, байесовские теоремы на основе искусственного интеллекта, алгоритмы LVQ (learning vector quantization) и схема фильтрации для систем на основе нейронных сетей. Рассмотрено прямое построение эффективной нейросетевой модели фильтрации спама, осуществляемой с использованием технологии распознавания баз данных. Приведены параметры доступа к нейронной сети как включать предопределенные статистические и нестатистические атрибуты сообщений. Рассмотрена структура нейросетевой технологии классификации электронных писем. Описана процедура анализа поступающих данных с использованием инструмента, входящего в аналитическую платформу Deductor Studio 5.3, которое в результате получается в виде тренировочного комплекта, пригодного для использования.*
Ключевые слова: *спам, искусственный интеллект, COM, алгоритм LVQ, информационная безопасность.*

Introduction

Today, information technology is evolving, and everyone in the community has become part of this technology. To use information technology, each user must have their own unique space in the infrastructure of the information system. This unique field is designed so that the user can identify himself in an additional information system and interact with other information systems. This unique

field is **email**. Users have the opportunity to register for all types of information systems through email addresses. Malicious users can access another user's email address by entering their username and password. The most common way to access your username and password is to send **spam** to this email.

With the growing number of Internet users, the task of protecting information has grown by one point - the problem of dividing emails into spam and non-spam categories. xcessive junk e-mail violates the user's access to the information resources that are required by the user, since they use the channel's capabilities [1]. In addition, they can be compromised by the loss of information in the event of loss of e-mail during filtering by a person or a special program. In addition to spam, malware can be sent that can completely or partially destroy or corrupt data.

Main part.

A number of malicious programs can be used to steal personal information in the system. For example, it can be used to steal users 'personal credit cards, steal their corporate names and passwords to access corporate bank accounts, and much more. Modern methods of protection against spam based on **linguistic signals** (detection of spam messages by inserting a predefined spam into the database), **message filtering rules** (various filtering methods set by the system administrator) are ineffective [1-2]. This is due to the need to attract more anti-spam experts to update signatures and policies. This is why modern linguistic signaling methods, such as spam filtering and messaging, have lost their relevance.

To solve this problem, it is proposed to use an approach based on the methods of artificial intelligence, especially artificial neural network. This approach requires training the classifier (classes that classify spam messages), marking important messages, setting model parameters and preparing training samples to evaluate the classification accuracy.

Thus, modern methods of combating spam require constant analysis of a person's message, and the system itself cannot develop these rules, that is, it cannot independently study it. If we consider a person a means of protection against spam, we can say that he or she has the ability to identify spam tags based on their experience and preferences, voluntary news and advertising subscriptions, but people can identify spam messages in certain templates will not fall. Thus, the task of creating a means of protection against spam is to have human-specific skills, such as *the ability to learn, a system of privileges and exceptions, contextual analysis and decision making.*

Information security, including spam protection, should be a process that requires non-trivial resources, such as human, machine and software, tripartite. One of the most promising areas for automation and improving the efficiency of information protection in order to free human resources is the introduction of neural network technologies in security systems. For example, ***systems for detecting and preventing network attacks using neural networks are common.*** [3] In these systems, neural networks analyze various network parameters (server response time, packet deviation from RFC standards, etc.), detect abnormal behavior and detect attacks that are not in the databases, due to the ability to summarize and study neural networks.

The spam filtering system is similar to a system for detecting and preventing attacks, and not only for network attacks, it detects spam messages in email content. The difference is that the proposed spam filtering system, **like the attack detection system, works on the OSI model**, and not on the network level, which checks the packets to be analyzed, not the flags, but the contents of the message itself.

There are many words and phrases that are often found in spam. However, there is no sufficient reason to classify such a message as spam. In this case, the individual will additionally pay attention to the content and meaning of the message, its general orientation, as well as the spelling, syntactic and morphological features of the message. Based on this package, you can decide whether this message is spam or not. Therefore, as part of a spam protection tool, neural network access parameters should include predefined statistical and non-statistical message attributes. These signs are:

Statistics Parameters:

- the exact number of words in the message, which should be spam;
- the exact number of phrases and phrases in the message that are suspected of spam.

Non-Statistical Data Parameters:

- semantic adjectives;
- text direction;
- morphological features - the correct expression of sentences and the relationship between parts of speech (Beckus-Naur method) [4] ;
- spelling - correct spelling, spelling replacement (for example, by typing "O" instead of "0" to reject the signal or vice versa).

Since the neural network works with numerical values, it is necessary to create a vector of numerical input of values from the above properties. [5]

A special dictionary containing spam-specific words is used to obtain statistical descriptions. The original message searches and counts words according to this dictionary. The most frequently used phrases are taken into account to increase the accuracy of decision making. This reduces the likelihood of false positives.

Analysis of the statistical characteristics of a neural network is similar to the Bayesian spam filter, in which you can specify the spam ratio for each word or phrase. However, unlike the Bayesian filter, the coefficients here are synaptic connections (weights) between neurons, which allow neural networks to efficiently identify new and previously unknown spam due to their ability to generalize accumulated experience. Thus, the external neural network is similar to the Bayesian filter, but they differ in internal architecture, additional functions and features of the neural network, which means that information in the neural network is presented in the form of numerical values, as well as semantic, phonetic and spelling values. Based on this, you can evaluate the text as spam, complement each other and evaluate it on the basis of many different parameters determined during the decision-making process.

This neural network is even more efficient when using the Coenon algorithm. Cohenon developed the SOM algorithm (SOM is a self-organizing map). SOM is a neural network that learns without a teacher, acting as visualization and clustering [7].

In the SOM algorithm, it is more difficult to categorize messages into spam or non-spam types. Since messages are grouped into 4 groups in this algorithm, some of these groups can be combined with each other, which means that when combining messages with spam and without spam, a gray area will appear in the algorithm and these zones will cause the problem of sorting into one type of message. To avoid this problem, we use the LVQ algorithm (LVQ - training vector quantization) together with the SOM algorithm [7].

Thus, the spam filtering algorithm based on the SOM algorithm described above can be supplemented by the steps of the LVQ algorithm. The neural network-based LVQ algorithm allows you to combine mail into specific classes in specific groups (for example, two classes: spam and non-spam messages, or four classes: spam, news, user ads, and simple messages). In addition, it is very difficult to directly determine the number of clusters in the SOM algorithm, since it does not include predefined definitions in the algorithm and requires careful selection of the coefficients used in the SOM algorithm. After setting up the first SOM-based neural network using the LVQ algorithm, all functions received from incoming e-mail messages and as vectors give you the opportunity to classify the message and decide whether to publish the message in spam or non-spam correspondence.

Since we are developing a scheme that blocks the classification of spam messages using these algorithms, it looks as follows. Figure 1 shows a generalized scheme of neural network technology for classifying emails into the spam and non-spam categories [7].

As can be seen from this picture, the use of neural technologies for the classification of emails includes the following key steps.

1. The choice of the structure of the neural network (determining the input and output parameters of the network, the number of layers and the number of layers in each layer);

2. Read (study) data from the electronic database of the selected type of neural network;

3. Use effective neural networks to classify new mail as spam or non-spam.

The direct construction of an effective neural network spam filtering model is carried out using database recognition technology, which includes the following steps:

4. Receive original emails, including examples of spam and non-spam;

5. Formation of a training model for preprocessing and training of neural networks;

6. Development of the structure of the neural network: the number of inputs, outputs, network levels and neurons in each layer;

7. Neural network training to create a spam filtering model;

8. Testing and evaluation of the neuralband spam filtering model.

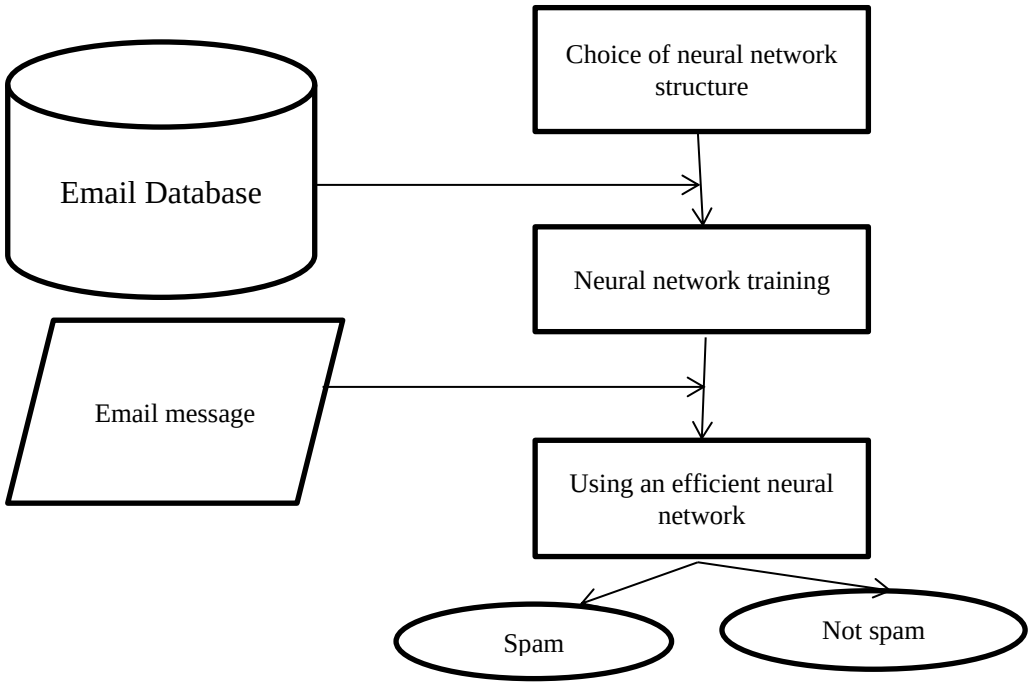

Figure 1. ***Structure of neural network technology for classifying emails.***

First, we need to select the most important parameters for analysis from text data. In other words, a specific number of parameters should be developed that characterize emails and allow them to be classified as spam or non-spam. The values of the selected parameters should be included in neural networks.

Among the many email markers, we single out the most important information labels needed to classify messages into spam or non-spam categories:

- frequency of capitalization;
- frequency of letter formation in letters;
- the number of different colors in the text;
- the size of the text;
- number of blank lines in the text.

To accomplish this task it is necessary to create data sets from different sources. The obtained source data is presented in the form of a table, where each row corresponds to a separate message and the message attribute of each column. Table cells contain values for functions that characterize a particular email message.

We used the tools included in the analytical platform Deductor Studio 5.3 to perform the described procedures for the analysis of incoming data [9 - 10]. After completing the procedures described, we get a training kit that is suitable for use, as shown in table 1.

Table 1.

P_1	P_2	P_3	P_4	P_5	TYPE
0	0	1	57	2	0
0	0	1	334	2	0
0	0.001	1	3	2	0
0	0.008	1	2	2	0
0	0.01	1	6	2	0
0	0.01	1	334	2	0
0	0.02	1	4	2	0
0.01	0.032	2	31	2	0
0.4	0	1	2	0	1
0	0	3	2	0	1
0	0	2	2	0	1
0.04	0	2	2	0	1
0.14	0	2	2	0	1
0.25	0	2	2	0	1
0.07	0	2	2	0	1
0.08	0	2	2	0	1

P_1- the frequency in which the capital letters appear in the email column, P_2 - frequency of occurrence of numbers in the message, P_3 – the number of different colors in the message text, P_4 – the message size in kilobytes, P_5 - the number of blank lines in the message. The last column of the TYPE Training Choice specifies the message type (1 - "spam", 0 - "not spam") [10].

The formulas used to calculate the e-mail parameter values are shown below:

- frequency of capital letters in the message text:

-

$$P_1 = \frac{n_1}{N_1},$$

where n_1 is the number of words in capital letters, and N_1 is the total number of words in the message;

- frequency of numbers appearing in the message:

$$P_2 = \frac{n_2}{N_2},$$

where n_2 is the number of messages in the message, N_2 is the total number of characters in the message;

- the number of different colors in the message text:

$$P_3 = N_3,$$

where N_3 - the number of different colors in the message;

- messages in kilobyte

$$P_4 = \frac{N_2}{1024}.$$

- the number of blank lines in the message

$$P_5 = N_4,$$

where N_4- the number of blank lines in the message.

We use the Neuroset utility, which is part of the Deductor Studio 5.3 analytic platform, to create a neural network model with email filtering. We define the inputs of the neural network from the fieldsP_1, P_2, P_3, P_4, P_5 Network output will be the only type of field.

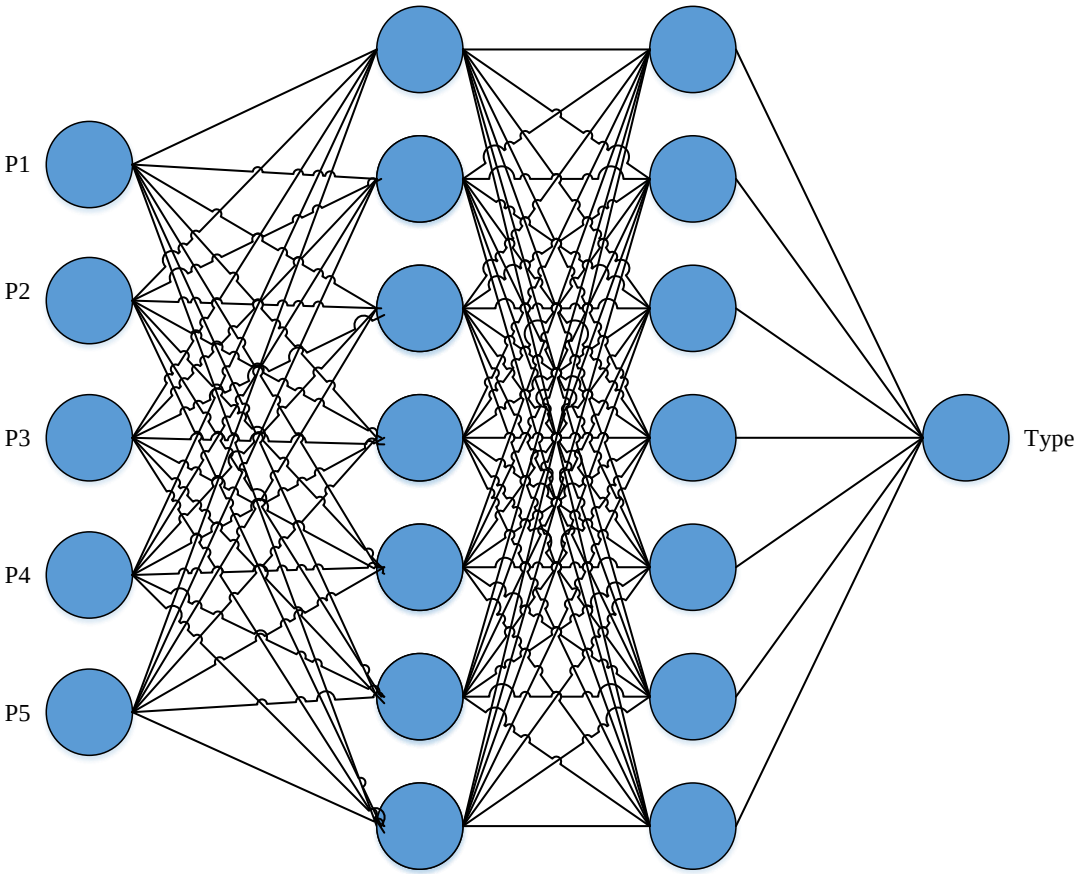

Figure 2. ***Neural network structure.***

As you can see in the figure above, all input parameters are analyzed in neural networks when a message is sorted into a single message, be it spam or spam. If the output is 0, the message will be spam, if 1 message is spam.

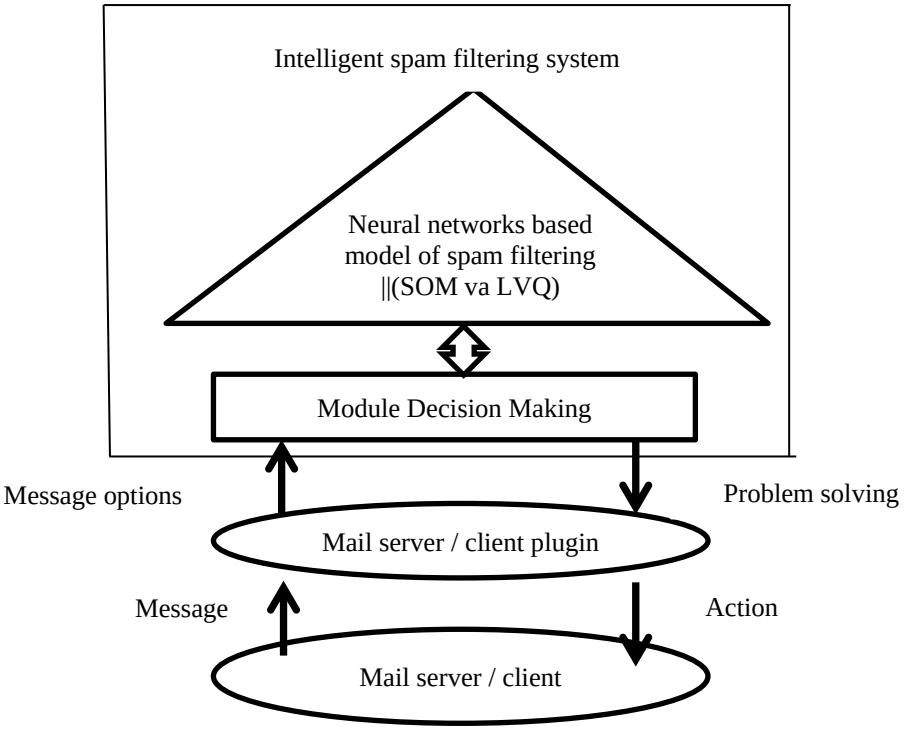

Figure 3. ***Schema of a neural network-based spam filtering system.***

As can be seen from the figure, the plug-in of the mail server (client) receives an email and selects the values of its parameters in accordance with the above formulas. The generated image of the mail message is transferred to the input of the decision-making module, which is directly connected to the neural network model of spam filtering. As a result, the problem of classifying email into spam or non-spam categories has been resolved [11].

Conclusion

Thus, we can analyze the problem of spam filtering and spam signaling in order to solve this problem, Bayesian theorems based on artificial intelligence and LVQ (vector quantization training LVQ) algorithms, and the structure of neural network technology for email classification. and a neural network spam filtering scheme. This proposed scheme requires training of classifiers used in the analysis of spam messages, labeling of important messages, setting model parameters and classification accuracy. In addition, the use of database recognition technology allows you to create effective neural networks to filter spam. For this, training systems, modeling of neural networks, their adequacy and classification are evaluated. It was shown that the neural network model can be effectively used as part of intelligent systems for filtering incoming mail. The diagram shown in Figure 3 has some actions for filtering email on the mail server (client). The combined use of SOM and LVQ algorithms further improves the accuracy of the results. If the email contains an incorrect classification status, the correct classification information can be found by accessing the database, and the system sets up a neural network to describe a new input or to create a new neural network model. system training in an updated sample. This type of spam protection is effective and can reduce a person's presence.

References:

1. Grey, A. et al. We read spam a lot: Prospective cohort study of unsolicited and unwanted academic invitations. BMJ 355, i5383 (2016).
2. Mazzarello S, Fralick M, Clemons M. A simple approach for eliminating spam. Curr Oncol2015;23:75-6. doi:10.3747/co.23.2860 pmid:26966417.
3. Bekmuratov Tulkun, Botirov Fayzullajon, Analysis of Integrated Neural Network Attack Detection System and User Behavior Models, // 2019 International Conference on Information Science and Communications Technologies (ICISCT), https://ieeexplore.ieee.org/document/9011869. 2019.
4. Kozak M, Iefremova O, Hartley J. Spamming in scholarly publishing: A case study.J Assoc Info Sci Tech, 2015, 10.1002/asi.23521.
5. Wang, W., Zhou, D. A multi-level approach to highly efficient recognition of Chinese spam short messages. Front. Comput. Sci. 12, 135–145 (2018).
6. MOHAMMED,M. A., GUNASEKARAN, S. S., MOSTAFA, S. A., MUSTAFA, A., & GHANI, M. K. A., Implementing an Agent-based Multi-Natural Language Anti-Spam Model. In 2018 International Symposium on Agent, Multi-Agent Systems and Robotics (ISAMSR) (pp. 1-5). IEEE. 2018, August
7. SHAFI'I, M. A., LATIFF, M. S. A., CHIROMA, H., OSHO, O., ABDUL-SALAAM, G., Abubakar, A. I., & Herawan, T., A review on mobile SMS spam filtering techniques. IEEE, 2017.
8. MIRZA, N., PATIL, B., MIRZA, T., & AUTI, R., Evaluating efficiency of classifier for email spam detec-tor using hybrid feature selection approaches. In Intelligent Computing and Control Systems (ICICCS), 2017 International Conference on (pp. 735-740). IEEE, 2017, June.
9. SINGH, M., Classification of Spam Email Using Intelligent Water Drops Algorithm with Naïve Bayes Classifier. In Progress in advanced computing & Intelligent Engineering (pp. 133-138). Springer, Singapore, 2019.
10. PAGANI, F., De ASTIS, M., GRAZIANO, M., LANZI, A., & BALZAROTTI, D., Measuring the Role of Grey-listing and Nolisting in Fighting Spam. In Dependable Systems and Networks (DSN), 2016 46th Annual IEEE/IFIP International Conference on (pp. 562-571), 2016, June.
11. FERRARA, E., Measuring social spam and the effect of bots on information diffusion in social media. In Complex Spreading Phenomena in Social Systems (pp. 229-255). Springer, Cham., 2018.