УДК 004.056.5
DOI: 10.15827/0236-235X.130.229-235
Дата подачи статьи: 02.03.20
2020. Т. 33. № 2. С. 229–235

# *Разработка метода защиты геоинформационных систем и пространственных данных на основе нейронной сети*

*Т.М. Татарникова* [1,2], *д.т.н., доцент, директор, tm-tatarn@yandex.ru*
*С.Ю. Степанов* [1], *к.т.н., доцент, stepanov.sergey.y@gmail.com*
*Я.А. Петров* [1], *к.т.н., доцент, yaroslav.petrov025@gmail.com*
*А.Ю. Сидоренко* [1], *старший преподаватель, sidorenko.ref@gmail.com*

[1] *Российский государственный гидрометеорологический университет, г. Санкт-Петербург, 192007, Россия*
[2] *Институт информационных систем и геотехнологий, г. Санкт-Петербург, 195196, Россия*

Для эффективного решения научно-практических и теоретических задач инвентаризации, анализа, моделирования, прогнозирования, управления системами окружающей среды и территориальной организацией сообществ широко используются геоинформационные системы.

Актуальность работы обусловлена необходимостью совершенствования методического аппарата для обнаружения возможных угроз в условиях динамики их роста и изменения концепций воздействий на пространственные данные в геоинформационных системах поддержки принятия решений. В ходе работы выполнен анализ требований к структуре систем защиты при обработке информации в геоинформационных системах. Приведен метод решения задач по созданию и поддержке эксплуатации систем защиты пространственной информации в геоинформационных системах.

Для решения задачи выбран и модифицирован алгоритм искусственной нейронной сети для обнаружения распределенных атак типа DDoS, целью которых является отказ в обслуживании и препятствование доступу легитимных пользователей к атакуемому приложению. На основе предложенного алгоритма разработана программа на языке высокого уровня Python. Программа включает в себя ряд компонентов, отвечающих за пополнение базы знаний нейронной сети, что, в свою очередь, позволяет строить произвольные архитектуры искусственной нейронной сети; анализатор трафика пакетов, так называемый сниффер, обеспечивающий фильтрацию пакетов по определенным сетевым протоколам модели взаимодействия открытых систем; связующий модуль, позволяющий направлять данные сниффера в базу знаний искусственной нейронной сети. Нейронная сеть может работать в двух режимах обучения: без учителя (самообучаемая) и с учителем, что, в свою очередь, дает пользователю возможность задавать начальные веса либо загружать файл с готовой базой знаний.

Результаты работы показывают, что искусственная нейронная сеть является одним из механизмов обнаружения потенциально опасных угроз в геоинформационных системах для поддержки принятия управленческих решений.

***Ключевые слова:*** *информационная безопасность, пространственные данные, защита данных, геоинформационная система (ГИС), нейронные сети.*

Среди многообразия информационных систем *геоинформационные системы* (ГИС) занимают особое место, поскольку в них к функциональным данным о текущих характеристиках и свойствах объектов реального мира добавляются пространственные данные, что позволяет найти местоположение интересующего объекта, отследить его передвижение, оценить окружающую обстановку и т.п. Развитие цифровых технологий показывает, что практически любая информационная система уже становится или станет в недалеком будущем ГИС. Использование геоинформации расширяет горизонты анализа при принятии решений и позволяет поднять решение прикладных задач по приоритетным направлениям на качественно новый уровень [1].

ГИС является аппаратно-программным, человеко-машинным (автоматизированным) комплексом, обеспечивающим актуализацию данных благодаря функциям сбора, обработки, отображения, передачи координируемых по пространству и времени данных, а также их интеграцию в представление о территории. Эти функции позволяют решать актуальные научно-практические и теоретические задачи анализа, моделирования, прогнозирования, управления системами окружающей среды и территориальной организацией сообществ и другие.

Популярность приложений ГИС делает их уязвимыми к DDoS-атакам, целью которых является отказ в обслуживании и препятствование доступу легитимных пользователей к атакуемому приложению. По открытым данным Лаборатории Касперского в целом наблюдается динамика роста DDoS-атак по сравнению с другими атаками (рис. 1). Это можно объяснить развитием новых интеллектуальных методов обеспечения информационной безопасности, что со стороны злоумышленника приводит к появлению технически более сложных (умных) атак.

Актуальность данной работы обусловлена необходимостью совершенствования методического аппарата для обнаружения возможных угроз в условиях динамики их роста и изменения концепций воздействий на пространственные данные в ГИС поддержки принятия решений [2].

### Описание предлагаемого решения

Для решения этой задачи предлагается использовать алгоритм и реализацию *искусственной нейронной сети* (ИНС) для обнаружения распределенных атак типа DDoS.

За основу взята обобщенная математическая модель искусственного нейрона [3, 4]: $out = \phi\left(\sum_{i=1}^{n} x_i w_i\right)$, где $out$ – выход нейрона; $\phi$ – функция активации; $\sum_{i=1}^{n} x_i w_i$ – взвешенная сумма.

В качестве функции активации была выбрана сигмоидальная функция: $\phi(u_k + b_k) = \frac{1}{1+\exp\left(-\alpha\left(u_k + b_k\right)\right)}$, где $\alpha$ – параметр наклона сигмоиды; $b_k$ – порог; $\phi(u_k + b_k)$ – функция активации; $u_k$ – взвешенная сумма.

Парадигма обучения нейронной сети – обучение с учителем. Идея состоит в том, что веса меняются согласно локальному градиенту функции ошибки [5].

Ошибка обучения на $n$-й итерации оценивается как $e_j(n) = d_j(n) - y_j(n)$, где $e_j(n)$ – сигнал ошибки нейрона $j$ на итерации $n$; $d_j(n)$ – желаемый отклик нейрона $j$; $y_j(n)$ – сигнал, генерируемый на выходе нейрона $j$.

Целевая функция обучения – ошибка результирующего вектора значений нейронной сети. В работе ошибка оценивалась в виде метрики MSE (Mean Squared Error) – среднеквадратической ошибки по обучающей выборке: $MSE = \frac{\sum_{i=1}^{N}\left(\vec{e} - \tilde{e}\right)^2}{N}$, где $N$ – количество итераций обучения нейронной сети; $\tilde{e}$ – результирующий вектор ожидаемых значений; $\vec{e}$ – результирующий вектор получаемых значений.

Таким образом, обучение нейронной сети – это автоматическая корректировка весов на значение: $\Delta w_{ji} = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n)$, где $\alpha$ – постоянная момента; $\eta$ – скорость обучения; $i$ – синаптическая связь нейрона $j$ с нейроном $i$; $\delta_j(n)$ – локальный градиент нейрона $j$; $y_i(n)$ – сигнал от нейрона $i$ в нейрон $j$.

Алгоритм процесса обучения ИНС приведен на рисунке 2.

В качестве примеров обучения включены шесть разных видов DDoS-атак: back, land,

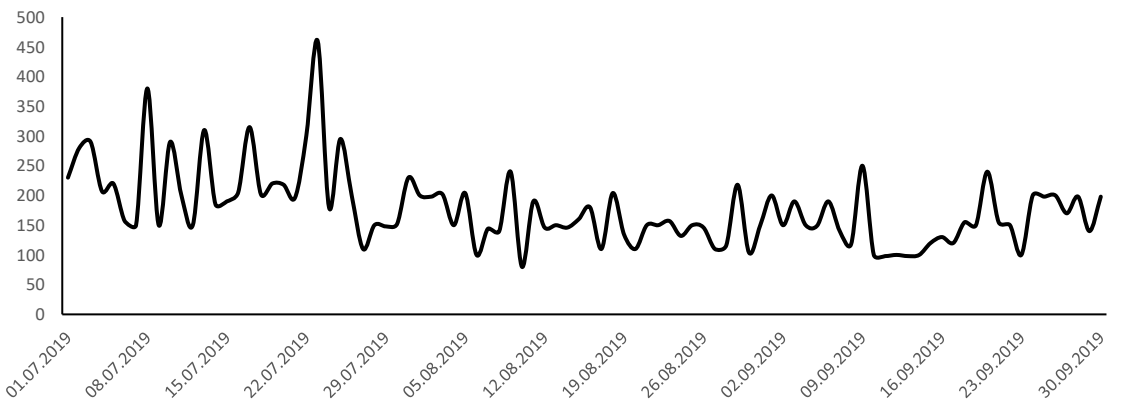


*Рис. 1. Динамика числа DDoS-атак в 3-м квартале 2019 г.*

*Fig. 1. DDoS attack number dynamics in the 3rd quarter of 2019*

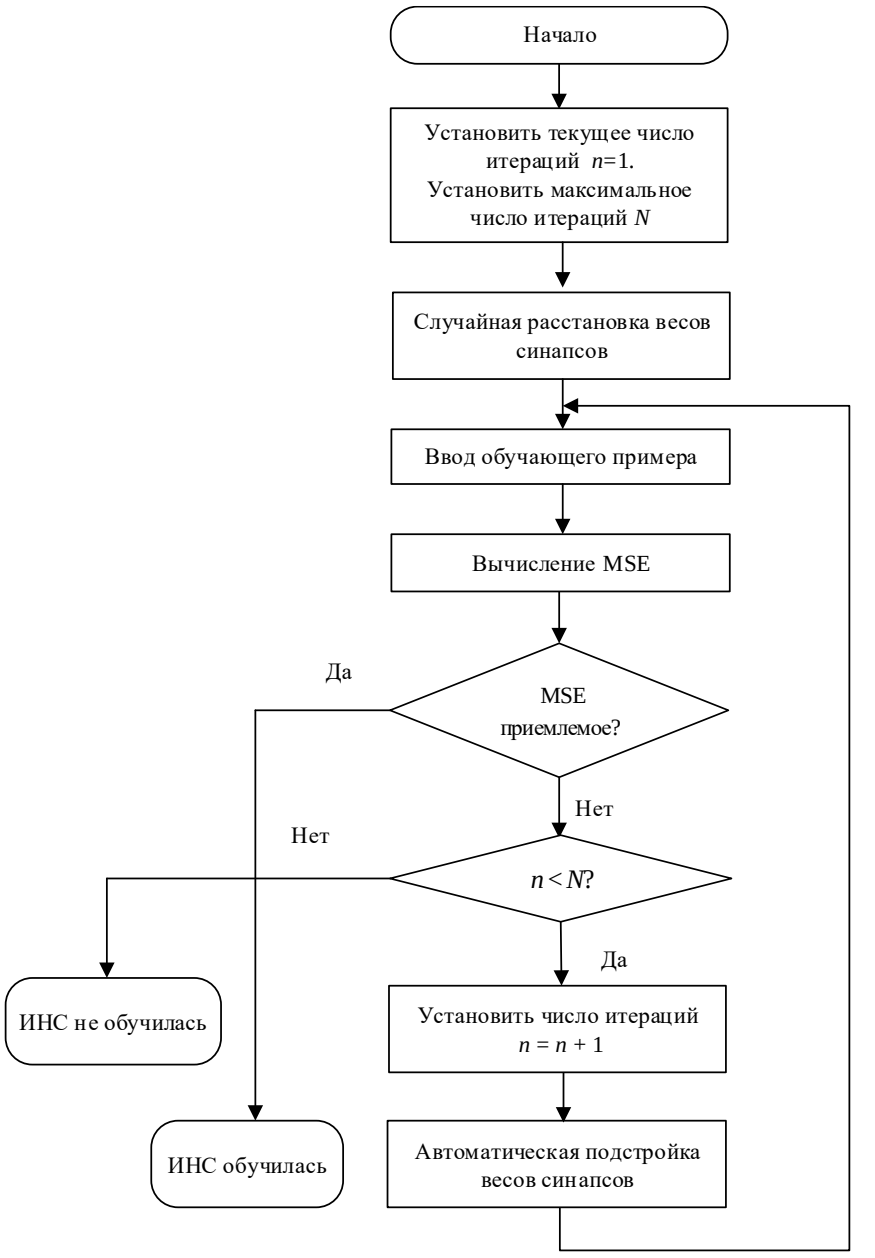


*Рис. 2. Блок-схема алгоритма обучения ИНС*

*Fig. 2. Block diagram of the ANN training algorithm*

neptune, pod, smurf, teardrop [6], каждая из которой содержит в себе различные параметры сетевого трафика (время соединения, службы, количество переданных байт, состояние порта, ошибочные пакеты, число пакетов с флагом URG и другие).

Для реализации нейронной сети выбран метод многослойного персептрона (рис. 3), архитектура которого в общем виде задается множеством параметров $A$: $A = \{H, N_I, N_H, N_O\}$, где $H$ – число скрытых слоев; $N_I$ – размер входного слоя; $N_j$ – размер $j$-го скрытого слоя, $j = \overline{1, H}$; $N_O$ – размер выходного слоя [7, 8].

В ходе эксперимента для обучения и тестирования нейронной сети была использована БД атак NSL-KDD [9, 10], каждая запись в ней представляет собой шаблон процесса передачи данных от IP-адреса источника к IP-адресу получателя по определенному протоколу. Шаблон содержит значения параметров и маркировку процесса – атака или не атака. Таким образом, данные, подаваемые на вход ИНС,

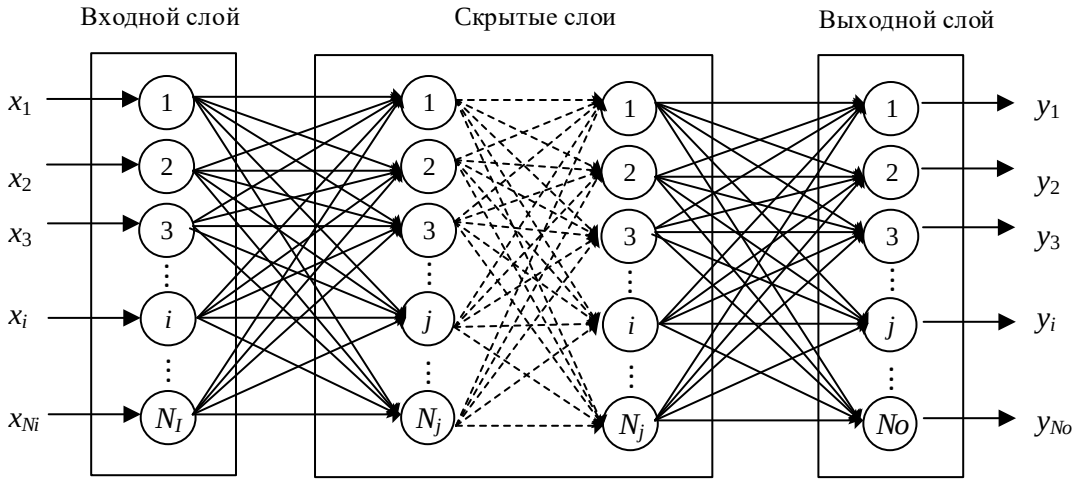


*Рис. 3. Обобщенная схема нейронной сети на основе многослойного персептрона*

*Fig. 3. Generalized scheme of a neural network based on a multilayer perceptron*

представляют собой текстовый файл, содержащий векторы как нормальной, так и аномальной активности. В БД атак NSL-KDD представлено 956 атак типа back, 18 – типа land, 41214 – типа neptune, 201 – типа pod, 2646 – типа smurf и 892 типа teardrop. В таблице 1 дано краткое описание некоторых параметров, позволяющих идентифицировать ту или иную атаку.

Таким образом, $N_I = 10$, $N_O = 6$, а $H$ и $N_j$ определены экспериментальным путем.

На основе описанного алгоритма разработана программа на языке Python. Программа включает в себя ряд компонентов, отвечающих за пополнение базы знаний нейронной сети, что, в свою очередь, позволяет строить произвольные архитектуры ИНС; анализ трафика пакетов (модуль сниффер), обеспечивающий фиксацию значений параметров пакетов; парсинг файла БД на векторы (записи) со значениями активности, позволяющий направлять дан-

*Таблица 1*

**Параметры из БД NSL-KDD**

*Table 1*

**Parameters from the NSL-KDD database**

| № | Параметр | Описание |
|---|---|---|
| 1 | count | Количество подключений к одному и тому же хосту назначения за последние две секунды |
| 2 | dst_host_srv_serror_rate | Процент соединений с SYN-ошибками при соединении по службе из dst_host_srv_count |
| 3 | srv_count | Число соединений с одной и той же службой за последние две секунды |
| 4 | serror_rate | Процент соединений с хостом из count с SYN-ошибками |
| 5 | srv_serror_rate | Процент соединений с SYN-ошибками при соединении по службе из srv_count |
| 6 | dst_host_count | Число соединений с тем же самым ip-адресом хоста назначения |
| 7 | dst_host_srv_count | Число соединений с тем же самым номером порта |
| 8 | dst_host_diff_srv_rate | Процент соединений по разным службам во время соединения по ip из dst_host_count |
| 9 | dst_host_same_src_port_rate | Процент соединений к тому же самому хосту-приемнику во время соединения по порту из dst_host_srv_count |
| 10 | dst_host_serror_rate | Процент соединений с хостом из dst_host_count с SYN-ошибками |

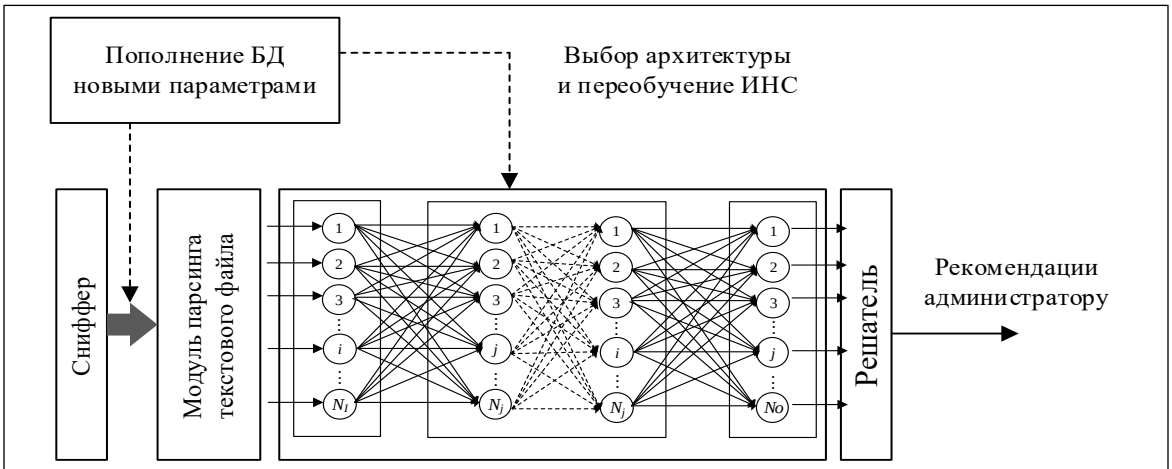


*Рис. 4. Структура программы по выявлению атак*

*Fig. 4. The program structure for detecting attacks*

ные сниффера на вход ИНС и принятие решений о возможных действиях администратора по противодействию выявленной атаке (рис. 4).

**Анализ результатов**

На рисунке 5 приведена диаграмма зависимости средней MSE от *H*, которая показывает, что у нейронной сети с одним скрытым слоем конечная MSE ниже, чем у сети с большим количеством слоев. Анализ зависимости указывает на выбор нейронной сети с одним скрытым слоем.

На рисунке 6 приведена диаграмма зависимости средней MSE от $N_H$, из которой видно, что конечная MSE приблизительно одинакова на всем разнообразии заданных в условии эксперимента значений $N_H$. В таком случае следует обратиться к значениям минимальной MSE. Результаты второго эксперимента указывают на выбор $N_H \in [6, 14]$.

В результате эксперимента была построена нейронная сеть со следующей архитектурой: 10 нейронов входного слоя, 1 скрытый слой, 9 нейронов скрытого слоя, 6 нейронов выходного слоя. Статистика по обучающей выборке представлена в таблице 2.

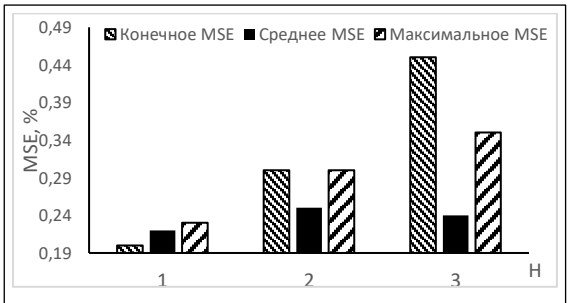


*Рис. 5. Диаграмма зависимости MSE от H*

*Fig. 5. The MSE dependent diagram from H*

*Таблица 2*

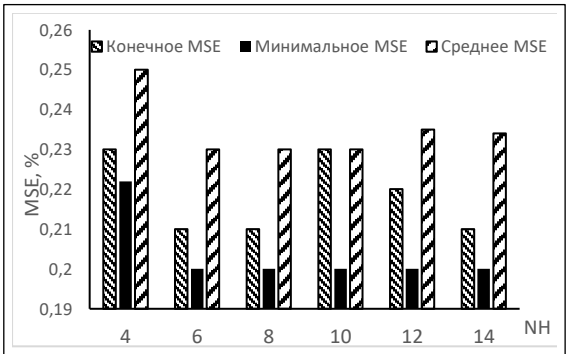


*Рис. 6. Диаграмма зависимости среднего значения MSE от $N_H$*

*Fig. 6. The dependent diagram of MSE average value from $N_H$*

**Обучающая выборка**

*Table 2*

**Training sample**

| Тип атаки | Количество атак |
|---|---|
| Отсутствие атаки | 3 500 |
| smurf | 1 500 |
| back | 400 |
| teardrop | 400 |
| neptune | 3 000 |
| Всего | 8 800 |

На графике снижения ошибки обнаружения аномального трафика (см. http://www.swsys.ru/uploaded/image/2020-2/2020-2-dop/21.jpg) видно, что число ошибок до оптимизации составляло 0,7 %, после 200 000 итераций ИНС была

оптимизирована до состояния 0,2 % ошибок. Исследование доказывает актуальность использования нейронной сети для обнаружения DDoS-атак и защиты данных в ГИС.

Стоит отметить, что при использовании нейронных сетей и других существующих методов защиты данных в совокупности в будущем возможно многократное повышение степени обнаружения потенциально опасных угроз в ГИС для поддержки принятия управленческих решений, что, в свою очередь, позволит построить эффективную и надежную систему обеспечения информационной безопасности данных.

## Заключение

Для выявления умных атак на ГИС-приложения предложено использовать технологию нейронных сетей. Предназначение нейронной сети – анализ трафика и принятие решения о том, является ли трафик нормальным или аномальным. Решение в виде рекомендаций администратору ГИС может не только информировать пользователя об инциденте или об атаке, но и указать точное время сессии и местоположение (пространственные данные) источника и получателя. Появление новой атаки связано с переобучением нейронной сети.

### *Литература*

1. Истомин Е.П., Сидоренко А.Ю., Колбина О.Н., Степанов С.Ю., Петров Я.А. Геоинформационная система управления пространственно-распределенными разнородными гидрометеорологическими данными для принятия управленческих решений по оптимизации регулирования отпуска тепла на ТЭЦ // Естественные и технические науки. 2019. № 4. С. 134–136.

2. Пилюгина К.Н. Применение нейронных сетей с целью обнаружения вторжений // Современные научные исследования и инновации. 2016. № 2. URL: http://web.snauka.ru/issues/2016/02/63248 (дата обращения: 02.02.2020).

3. Гелиг А.Х. Введение в математическую теорию обучаемых распознающих систем и нейронных сетей. М., 2017. 224 с.

4. Каллан Р. Нейронные сети. Краткий справочник. М.: Вильямс, 2017. 288 с.

5. Тархов Д.А. Нейросетевые модели и алгоритмы. М.: Радиотехника, 2017. 787 с.

6. Пальчевский Е.В., Христодуло О.И. Разработка метода самообучения импульсной нейронной сети для защиты от DDoS-атак // Программные продукты и системы. 2019. Т. 32. № 3. С. 419–432. DOI: 10.15827/0236-235X.127.419-432.

7. Татарникова Т.М. Анализ данных. СПб: Изд-во СПбГЭУ, 2018. 85 с.

8. Хайкин С. Нейронные сети: полный курс. М.: Диалектика, 2019. 1104 с.

9. Татарникова Т.М., Сидоренко А.Ю., Степанов С.Ю., Петров Я.А. Реализация метода для защиты пространственных данных ГИС на основе нейронной сети // Естественные и технические науки. 2019. № 1. С. 134–136.

10. Частикова В.А., Картамышев Д.А., Власов К.А. Нейросетевой метод защиты информации от DDoS-атак // Современные проблемы науки и образования. 2015. № 1. Ч. 1. С. 183–190.

Software & Systems
DOI: 10.15827/0236-235X.130.229-235

Received 02.03.20
2020, vol. 33, no. 2, pp. 229–235

## Development of a geoinformation systems protecting method and spatial data based on a neural network

***T.M. Tatarnikova*** [1,2], *Dr.Sc. (Engineering), Associate Professor, Director, tm-tatarn@yandex.ru*
***S.Yu. Stepanov*** [1], *Ph.D. (Engineering), Associate Professor, stepanov.sergey.y@gmail.com*
***Ya.A. Petrov*** [1], *Ph.D. (Engineering), Associate Professor, yaroslav.petrov025@gmail.com*
***A.Yu. Sidorenko*** [1], *Senior Lecturer, sidorenko.ref@gmail.com*

[1] *Russian State Hydrometeorological University, St. Petersburg, 192007, Russian Federation*
[2] *Institute of Information Systems and Geotechnologies, St. Petersburg, Russian Federation*

**Abstract.** The GIS usage is necessary for effective solution in scientific, practical and theoretical problems of inventory, analysis, modeling, forecasting, environmental system management, community territorial organization.

The relevance of this paper is due to the need to improve the methodological apparatus for detecting possible threats in the context of their growth dynamics and changing the impact concepts on spatial data in GIS for decision support. In the course of scientific paper, the requirements to the protection system structure, there is analyze during information processing in geographic information systems. The article presents a method for solving problems of creating and supporting the spatial information protection system operation in a GIS.

To solve this problem, the authors selected an artificial neural network algorithm and modified to detect distributed DDoS attacks, the purpose of which is «service denial» and prevent legitimate users from accessing the attacked application.

Based on the presented algorithm, the authors developed a program in a high-level language – Python. This program includes a component number responsible for: replenishing the knowledge base of a neural network, which in turn allows you to build arbitrary ANN architectures; packet traffic analyzer, the so-called sniffer, which provides packet filtering according to certain network protocols of the OSI model; a connecting module that allows you to send sniffer data to the ANN knowledge base. The resulting neural network can operate in two training modes: without teacher (self-taught), with teacher, which in turn allows the user to set the initial weights, or specify a file with a ready-made knowledge base.

**Keywords:** information security, spatial data, data protection, GIS, neural networks.

### *References*

1. Istomin E.P., Sidorenko A.Yu., Kolbina O.N., Stepanov S.Yu., Petrov Ya.A. Geoinformation system for managing spatially distributed heterogeneous hydrometeorological data for making managerial decisions to optimize the regulation of heat supply at thermal power plants. *Natural and Tech. Sci.*, 2019, no. 4, pp. 134–136 (in Russ.).
2. Piliugina K.N. Artificial neural network approaches to intrusion detection. *Modern Scientific Researches and Innovations*, 2016, no. 2. Available at: http://web.snauka.ru/issues/2016/02/63248 (accessed February 02, 2020) (in Russ.).
3. Gelig A.Kh. *Introduction to the Mathematical Theory of Learner Recognition Systems and Neural Networks*. Moscow, 2017, 224 p. (in Russ.).
4. Callan R. *Neural Networks. A Quick Reference*. Moscow, Williams, 2017, 288 p. (in Russ.).
5. Tarkhov D.A. *Neural Network Models and Algorithms*. Moscow, 2017, 787 p. (in Russ.).
6. Palchevsky E.V., Khristodulo O.I**.** Development of a spiking neural network with the possibility of high-speed training to neutralize DDoS attacks**.** *Software & Systems*, 2019, vol. 32, no. 3, pp. 419–432. DOI: 10.15827/0236-235X.127.419-432 (in Russ.).
7. Tatarnikova T.M. *Data Analysis*. St. Petersburg, 2018, 85 p. (in Russ.).
8. Khaikin S. *Neural Networks: Full Course*. Moscow, 2019, 1104 p. (in Russ.).
9. Tatarnikova T.M., Sidorenko A.Yu., Stepanov S.Yu., Petrov Ya.A. Implementation of a method for protecting spatial GIS data based on a neural network. *Natural and Tech. Sci*., 2019, no. 1, pp. 134–136 (in Russ.).
10. Chastikova V.A., Kartamyshev D.A., Vlasov K.A. The neural network method of protecting information from DDoS attacks. *Modern Problems of Science and Education*, 2015, no. 1, pt. 1, pp. 183–190 (in Russ.).

**Для цитирования**

Татарникова Т.М., Степанов С.Ю., Петров Я.А., Сидоренко А.Ю. Разработка метода защиты геоинформационных систем и пространственных данных на основе нейронной сети // Программные продукты и системы. 2020. Т. 33. № 2. С. 229–235. DOI: 10.15827/0236-235X. 130.229-235.

**For citation**

Tatarnikova T.M., Stepanov S.Yu., Petrov Ya.A., Sidorenko A.Yu. Development of a geoinformation systems protecting method and spatial data based on a neural network. *Software & Systems*, 2020, vol. 33, no. 2, pp. 229–235 (in Russ.). DOI: 10.15827/0236-235X.130.229-235.