УДК 004.056.5 и 004.8

КЛАССИФИКАЦИЯ СЕТЕВЫХ АТАК НА ОСНОВЕ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ С 1D-СВЕРТОЧНЫМИ И РЕКУРРЕНТНЫМИ СЛОЯМИ

О.С. Амосов
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: osa18@yandex.ru

Д.С. Магола
Комсомольский-на-Амуре государственный университет
Россия, 681013, Комсомольск-на-Амуре, ул. Ленина, 27
E-mail: dmagola@list.ru

Ф.Ф. Пащенко
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: pif-70@yandex.ru

С.Г. Амосова
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: amosovasg@yandex.ru

Ключевые слова: сетевая атака, классификация, глубокая нейронная сеть, сверточный слой, рекуррентный слой.

Аннотация: Представлена постановка задачи классификации сетевых атак и ее решение на основе технологии глубоких нейронных сетей с использованием одномерных сверточных и рекуррентных слоев. Приводятся результаты моделирования с использованием базы данных сетевых атак KDD.

1. Введение

Все больше работ отечественных и зарубежных авторов для решения разнообразных задач информационной безопасности (ИБ) используют различные сочетания моделей и методов искусственного интеллекта [1, 2, 4, 7, 8, 12-14]: экспертные системы, искусственные нейронные сети (НС), модели основанные на нечеткой логике, генетические алгоритмы, роевой интеллект, а также гибридные подходы с другими математическими направлениями. Так, авторами, в работе [2] был предложен алгоритм классификации атак на основе интеллектуальных технологий с элементами фрактального и вейвлет-анализа. Представленная здесь работа является продолжением исследований, изложенных в [2], однако в качестве базовой технологии для решения задачи классификации атак рассмотрены глубокие нейронные сети (ГНС) [5, 7, 9]. Прорыв в обуче-

нии НС и популяризации ГНС произошел после появления работы [10]. При этом такие задачи как задачи ИБ являются частным случаем задач обработки временных последовательностей.

2. Постановка задачи сетевой классификации

Под объектом будем понимать в дальнейшем сетевую атаку или аномалию.

Дано: множество Ω, в котором хранится описание объектов $\omega \in \Omega$, заданных признаками $x_i, i = \overline{1,n}$, совокупность которых для объекта ω представлена векторными описаниями $\mathbf{x} = \mathbf{\Phi}(\omega) = (x_1(\omega), x_2(\omega), \ldots, x_n(\omega))^T$; множество классов $\mathbf{B} = \{\beta_1, \beta_2, \ldots, \beta_c\}$, c –количество классов; априорная информация, которая представлена обучающим множеством $\mathbf{D} = \{(\mathbf{x}^j, \beta^j)\}$, $j = \overline{1,L}$, заданным таблицей, каждая строка j которой содержит векторное описание объекта $\mathbf{x} = \mathbf{\Phi}(\omega)$ и метку класса β_k, $k = \overline{1,c}$. Заметим, что обучающее множество характеризует неизвестное отображение * $\mathbf{F}: = \Omega \to \mathbf{B}$.

Требуется по поступающим фрагментам $\mathbf{I}_t$ непрерывного сетевого трафика $\mathbf{V} = (\mathbf{I}_1, \ldots, \mathbf{I}_t, \ldots, \mathbf{I}_\tau)$ и априорной информации, заданной обучающим множеством $\mathbf{D} = \{(\mathbf{x}^j, \beta^j)\}$, $j = \overline{1,L}$ для глубинного обучения НС с учителем, решить задачу распознавания объектов: обнаружить объекты ω в виде оценки признаков $\tilde{\mathbf{x}}$ с помощью нейронных сетей, реализующих отображение $\mathbf{F}_1: \mathbf{I}_t \to \tilde{\mathbf{x}}$, и классифицировать их с использованием отображения $\mathbf{F}_2: \tilde{\mathbf{x}} \to \beta_k$, $k = \overline{1,c}$ в соответствии с заданным критерием $\mathrm{P}(\tilde{\mathbf{x}})$, минимизирующим вероятность ошибки классификации.

Таким образом, необходимо найти отображение $\mathbf{F}: \mathbf{I}_t \to \beta_k$, $k = \overline{1,c}$, при котором $\mathbf{F}$– является набором функций и нейросетевых алгоритмов $\mathbf{f}_i$, $i = \overline{1,N_f}$.

Заметим, что к простейшей задаче сетевой безопасности сводится задача классификации с целью отнесения рассматриваемого временного ряда сетевого трафика к одному из двух классов {«отсутствие атаки», «присутствие атаки»}, $\beta^t \in \mathbf{B} = \{-1; +1\}$. Это задача бинарной классификации. Однако в задачах ИБ отнесение сетевого взаимодействия лишь на два класса {«отсутствие атаки», «присутствие атаки»} не является желаемым решением, поскольку факт наличия атаки без понимания ее природы не позволяет за адекватное время применить методы борьбы с атакой. Поэтому разумным является именно многоклассовая классификация.

3. Решение задачи сетевой классификации на основе ГНС с 1D-сверточными и реккурентными слоями

В общем виде решение задачи классификации сетевых атак с помощью ГНС можно представить с помощью схемы на рис. 1.

Для экспериментального моделирования использовалась общедоступная база данных (БД) KDD [15]. БД содержит почти 5 миллионов (4 891 469) классифицированных по 22 типам экземпляров сетевого состояния на основании 41 признака (продолжительность, протокол, сервис, количество байт и др.). Для получения результатов, описанных ниже, использовалось эталоны трех состояний сети: нормальное состояние и 2 атаки: Neptune, smurf. Это объясняется тем, что именно эталоны этих состояний составляют более 99,2% всей БД. Таким образом, для моделирования использовалась БД 3-х состояний сети, содержащая матрицу 4 853 584×42, где каждая запись определяет состоя-

ние компьютерной сети по 41-му признаку, а последний столбец – номер класса.

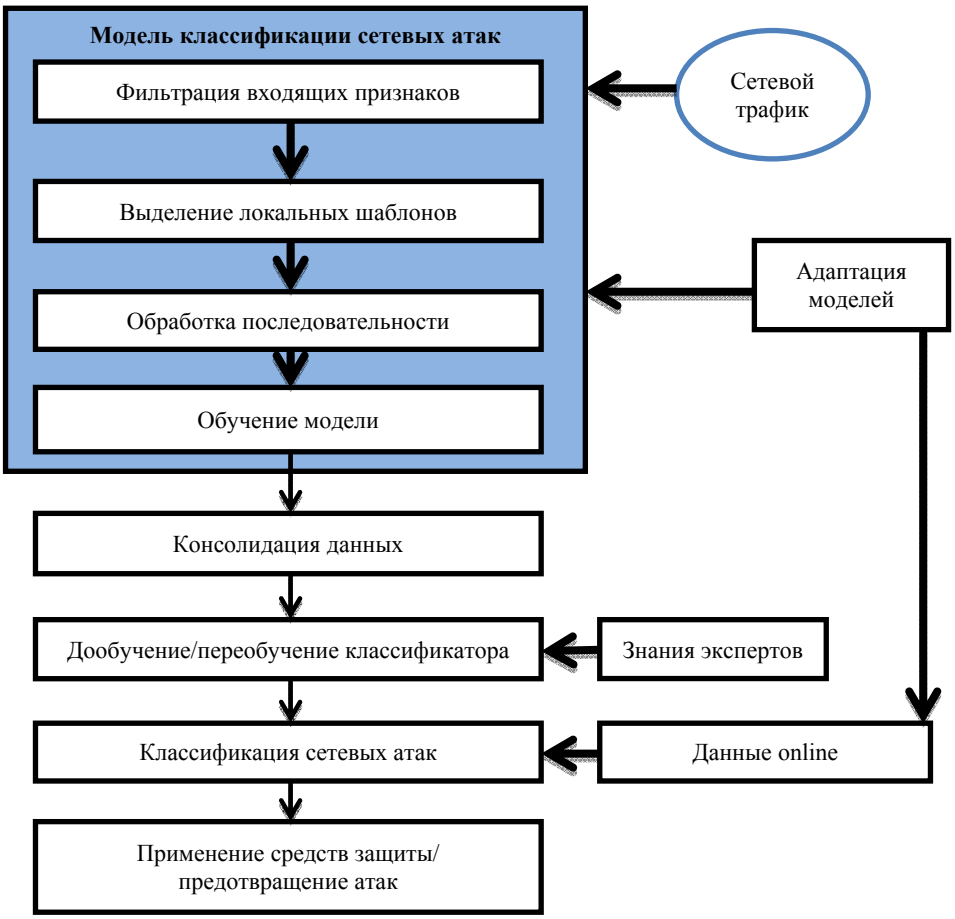

Рис. 1. Решение задачи классификации сетевых атак.

Моделирование производилось на ПК со следующими параметрами Intel Core i7-4710HQ, 4 ядра по 2,5 GHz, ОЗУ 6 Гб, 64-разрядная операционная система Windows.

3.1. LSTM-сеть

НС с рекуррентным слоем долгой краткосрочной памяти (LSTM-сеть) лишена проблемы затухания градиента и при этом является основой для предложенной позже GRU-сети [3, 6, 11]. Исходя из данных предпосылок, при моделировании использовалась LSTM-сеть (рис. 2).

Рассматривались различные варианты сочетаний объема обучающей и тестовой выборок. Моделирование производилось в системе MATLAB 2018. Сеть начинается с входного слоя последовательности, за которым следует слой LSTM. Чтобы спрогнозировать метки классов, сеть заканчивается полностью связанным слоем, слоем *softmax* и выходным слоем классификации. Результаты моделирования представлены в таблице 1. Для оптимизации машинных затрат и получения результата за меньшее время набор обучающих данных из двумерного вида преобразовывался в трехмерный.

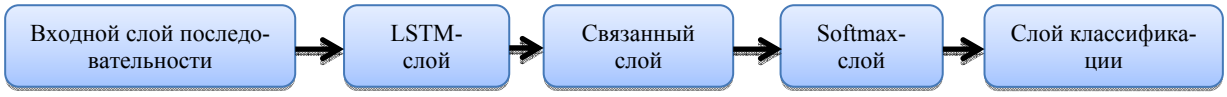

Рис. 2. Структура LSTM-сети.

Таблица 1. Результаты моделирования LSTM-сети.

№ эксперимента	Размер обучающей выборки	Размер тестовой выборки	Размер преобразованного тензора	Точность на обучающей выборке, %	Точность на тестовой выборке, %	Время обучения, мин
1	3 000 x 41	3 000 x 41	150 x 20 x 41	100	100	0,7
2	12 000 x 41	3 000 x 41	600 x 20 x 41	100	100	1,5
3	120 000 x 41	30 000 x 41	6 000 x 20 x 41	100	100	16
4	1 200 000 x 41	300 000 x 41	60 000 x 20 x 41	100	99,29	75
5	210 000 x 41	60 000 x 41	420 x 500 x 41	100	100	1,1
6	2 100 000 x 41	600 000 x 41	420 x 5 000 x 41	100	84,17	22
7	300 000 x 41	2 100 000 x 41	600 x 500 x 41	100	95,9	12

3.2. Одномерная сверточная сеть

Хорошие результаты сверточных сетей по распознаванию изображений позволяют использовать их для обработки временных последовательностей, поскольку время можно рассматривать как пространственное измерение, подобно высоте или ширине двумерного изображения. Использовалась одномерная сверточная сеть, состоящая из двух стопок сверточных и субдискретизирующих слоев с/без LSTM слоем (рис. 3). Моделирование производилось в среде Python 3.6, фреймворк Keras. В качестве функции активации сверточного слоя использовалась функция ReLU: $\sigma(x) = \max(0, x)$. Результаты моделирования представлены в таблице 2.

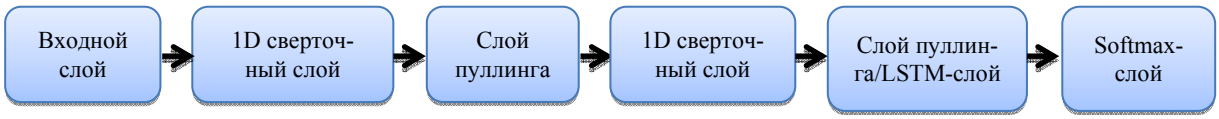

Рис. 3. Структура одномерной сверточной сети.

Из таблицы 2 видно, что использование одномерной сверточной сети дает хорошую точность на обучающих и тестовых выборках. 100% точность обучения достигнута была примерно за 1,5 часа на ПК с описанными выше характеристиками, при этом обучающая выборка содержала 45 000 примеров. Однако, в отличии от примера моделирования сети только с LSTM-слоем, в данном примере исходные данные не преобразовывались из двухмерного вектора в трехмерный, таким образом вероятность искажения исходной информации сведена к минимуму.

Таблица 2. Результаты моделирования сверточной сети.

№	Размер обучающей выборки	Размер тестовой выборки	Наличие рекуррентного слоя	Точность на обучающей выборке, %	Точность на тестовой выборке, %	Время обучения, мин
1	3 000 x 41	3 000 x 41	Нет	100	98,8	6
2	20 000 x 41	10 000 x 41	Нет	100	100	34
3	30 000 x 41	20 000 x 41	Нет	100	98	54
4	45 000 x 41	45 000 x 41	Да	100	100	83

4. Заключение

Дана постановка задачи сетевой классификации и рассмотрена возможность использования технологии глубоких нейронных сетей для решения задачи сетевой классификации.

Необходимость сочетать или не сочетать в одной модели возможности глубоких рекуррентных и сверточных сетей исследователь должен определять индивидуально, исходя из ресурсных ограничений, качества обучающего набора данных и специфики предметной области. Описанный выше подход, как и в целом принципы глубокого машинного обучения, ориентируется на обучение с учителем, предполагая, что имеется достаточно большая обучающая выборка (10 000 примеров и более). С другой стороны, при отсутствии таковой, что не исключается в реальной ситуации, выходом является использование нечетких моделей или моделирование на основе обучения с подкреплением.

Список литературы

1. Амосов О.С., Магола Д.С., Муллер Н.В. Фрактальный и вейвлет-анализ телекоммуникационных рядов информационной системы // Ученые записки КнАГТУ. 2016. № I-1 (25). С. 28-36.
2. Амосов О.С., Магола Д.С., Баена С.Г. Сетевая классификация атак в задачах информационной безопасности на основе интеллектуальных технологий, фрактального и вейвлет-анализа // Ученые записки КнАГТУ. 2017. № IV-1 (32). С. 19-29.
3. Будыльский Д.В. GRU и LSTM: современные рекуррентные нейронные сети // Молодой ученый. 2015. № 15. С. 51-54. URL https://moluch.ru/archive/95/21426/.
4. Корченко А.Г. Построение систем защиты информации на нечетких множествах. Теория и практические решения. К.: МК-Пресс, 2006. 320 с.
5. Шолле Ф. Глубокое обучение на Python. СПб: Питер, 2018. 400 с.
6. Chung Junyoung, Gulcehre Caglar, Cho KyungHyun, Bengio Yoshua. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. 2014. arXiv:1412.3555.
7. Dotcenko S., Vladyko A., Letenko I. A fuzzy logic-based information security management for software-defined networks // 16th International Conference on Advanced Communication Technology. Pyeongchang, 2014. P.. 167-171. doi: 10.1109/ICACT.2014.6778942.
8. Hamamoto A.H., Carvalho L.F., Proenca, M.L. ACO and GA metaheuristics for anomaly detection // 34th International Conference of the Chilean Computer Science Society (SCCC). Santiago, 2015. P. 1-6. doi: 10.1109/SCCC.2015.7416569.
9. Hinton G.E., Osindero S., Teh Y.-W. A fast learning algorithm for deep belief nets // Neural computation. 2006. Vol. 18, No. 7. P. 1527-1554.
10. Hinton G.E., Salakhutdinov R.R. Reducing the dimensionality of data with neural networks // Science. 2006. Vol. 313. P. 504-507.
11. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. 1997. Vol. 9, No. 8.
12. Katsupeev A.A., Shcherbakova E.A., Vorobyev S.P. Comparison of evolutionary algorithms used to solve the optimization problem of information security of distributed systems // 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). Chelyabinsk, 2016. P. 1-3. doi: 10.1109/ICIEAM.2016.7911608.
13. Otero A.R., Tejay G., Daniel Otero L., Ruiz-Torres A.J. A fuzzy logic-based information security control assessment for organizations // IEEE Conference on Open Systems. Kuala Lumpur, 2012. P. 1-6. doi: 10.1109/ICOS.2012.6417640.
14. Wu M.S. Genetic algorithm using discrete cosine transformation for fractal image encode // International Conference on Information Security and Intelligent Control. Yunlin, 2012. P. 309-312. doi: 10.1109/ISIC.2012.6449768.
15. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.