7. Шапкин А.В., Кубасов И.А., Иванов А.И. Развитие отечественного нейросетевого искусственного интеллекта в защищенном исполнении. Вестник Воронежского института ФСИН России. 2019. № 4. С. 132-144.
8. Кубасов И.А. Проблемные вопросы применения технологий искусственного интеллекта в деятельности органов внутренних дел Российской Федерации. Вестник Воронежского института МВД России. 2021. № 3. С. 180-186.

УДК 003.26+004.032.26

О ПОДХОДАХ К РЕШЕНИЮ ЗАДАЧ КРИПТОЛОГИИ НА ОСНОВЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ И МАШИННОГО ОБУЧЕНИЯ

М.В. МАЛЬЦЕВ, Ю.С. ХАРИН
Учреждение Белорусского государственного университета
«Научно-исследовательский институт прикладных
проблем математики и информатики»
Минск, 220030, Республика Беларусь

Введение. Машинное обучение (machine learning) – сравнительно молодое направление анализа данных, включающее в себя методы и алгоритмы, которые позволяют автоматически выявлять закономерности в данных и затем использовать обнаруженные закономерности для решения различных задач: прогнозирования, классификации и кластеризации, обнаружения аномалий и т.д. [1]. Одним из наиболее перспективных и активно развивающихся методов машинного обучения являются искусственные нейронные сети (ИНС), широкое использование которых началось в 2005–2006 годах, когда канадские ученые Джеффри Хинтон и Йошуа Бенджи научились обучать так называемые глубокие нейронные сети [2]. Применение глубоких ИНС позволило значительно улучшить качество распознавания речи и изображений, существенно продвинуться в ряде задач биоинформатики, создать системы искусственного интеллекта, способные на сопоставимом с человеком уровне управлять автомобилем и играть в интеллектуальные игры. Закономерно, что интерес к ИНС как к эффективному инструменту решения самых разнообразных задач появился и в криптографическом сообществе. Применению ИНС и других методов машинного обучения в криптологии посвящена настоящая статья.

1. Машинное обучение и искусственные нейронные сети. Машинное обучение предполагает, что программа обучается по мере накопления опыта относительно некоторого класса задач T и целевой функции P, если качество решения этих задач (относительно P) улучшается с получением нового опыта [3]. Выделяют два основных класса задач машинного обучения: обучение с учителем (supervised learning) и обучение без учителя (unsupervised learning). Обучение с учителем подразумевает наличие множества обучающих данных (часто называемых размеченными), для которых известно значение целевой функции P (например, набор фотографий с отмеченными на них объектами), и задача состоит в том, чтобы, обучившись на размеченных данных, сделать вывод относительно новых (тестовых) данных (в нашем примере – выделить объекты на неразмеченных фотографиях). В задачах

обучения с учителем используются такие методы как деревья решений, метод ближайших соседей, логистическая регрессия, нейронные сети. В обучении без учителя размеченные данные отсутствуют, и задача, как правило, состоит в кластеризации – разделении данных на классы (число классов может быть как заранее заданным, так и неизвестным) в зависимости от определенных характеристик. Примером такой задачи является распределение программой-агрегатором новостей по тематическим рубрикам. В задачах обучения без учителя используются статистические алгоритмы кластеризации, метод главных компонент, нейронные сети.

Таким образом, важным методом машинного обучения, используемым как в задачах с учителем, так и без учителя, являются ИНС. Основным структурным элементом ИНС является перцептрон или нейрон (см. рис. 1). Перцептрон получает вектор входных значений $X = (x_1, ..., x_n) \in \mathrm{R}^n$ и вектор весов $W = (w_1, ..., w_n) \in \mathrm{R}^n$, соответствующих этим входным значениям. Выход перцептрона $y = y(X, W)$ имеет вид:

$$y = h\left(\sum_{i=1}^{n} w_i x_i\right),$$

где h – функция активации, в качестве которой чаще всего используются логический сигмоид, гиперболический тангенс, функция Хевисайда, функция ReLU (rectified linear units). Основной задачей при работе с нейронной сетью является ее обучение –определение вектора W, минимизирующего потери при обработке входных значений из заданного множества (например, функция потерь может определятся как суммарное расстояние между истинными значениями $y_0(X)$ и выходными значениями перцептрона $y(X, W)$.

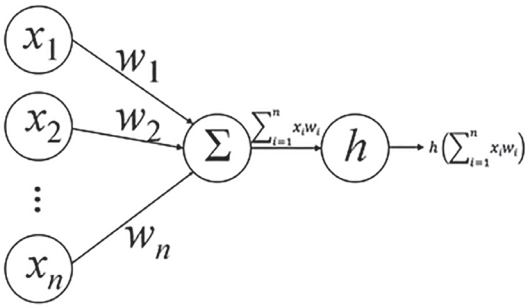

Рис. 1. Схема перцептрона

Возможности одного перцептрона сильно ограничены, поэтому их объединяют друг с другом (как правило, в слои) в нейронную сеть. На рисунке 2 приведена схема нейронной сети, в состав которой входят входной и выходной слои, а также m–2 скрытых слоя. В зависимости от строения (архитектуры) выделяют различные типы ИНС: сверточные сети, рекуррентные сети, генеративно-состязательные нейронные сети и другие [2]. Приведем далее кратное описание нескольких популярных архитектур ИНС.

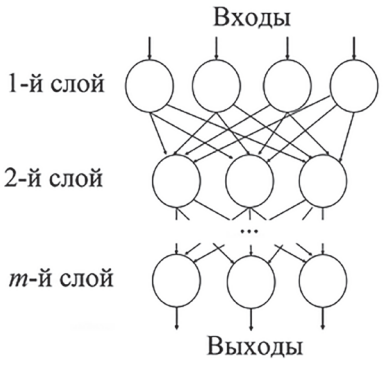

Рис. 2. Схема нейронной сети

Сверточные ИНС (convolutional neural network, CNN) предложены французским ученым в области искусственного интеллекта Яном ЛеКуном. Такие сети состоят из чередующихся слоев двух типов: сверточных (convolution) и субдискретизирующих (pooling). Сверточные ИНС широко применяются для распознавания изображений.

В рекуррентных ИНС (recurrent neural network, RNN) нейроны образуют направленную последовательность, вследствие чего такие сети могут обрабатывать серии событий во времени или последовательные пространственные данные. Такие сети используются для распознавания рукописного текста, распознавания речи.

Генеративно-состязательные нейронные сети (generative adversarial network, GAN), предложенны сотрудником компании Google Яном Гудфеллоу. В данной модели имеются две нейронные сети: генератор и дискриминатор. Генератор порождает объекты, принадлежащие различным классам, а дискриминатор пытается отличать объекты из разных классов. Применяются такие сети, например, для генерации и улучшения качества изображений.

2. Подходы к решению задач криптологии на основе искусственных нейронных сетей и машинного обучения. Идеи использовать машинное обучение в криптологии высказывались еще в начале 90-х годов Ривестом [4], но широкое распространение получили в последнее десятилетие вследствие развития методов обучения ИНС и роста вычислительных мощностей. ИНС, используя большие объемы информации, позволяют выявлять скрытые закономерности в данных сложной структуры, что способствует их применению в задачах криптоанализа [5–9]. Отметим, что в открытой печати, как правило, анализируются упрощенные версии криптографических алгоритмов: в работе [5] исследовалась упрощенная версия алгоритма DES, в работе [6] – алгоритм Speck32/64 с уменьшенным числом раундов. В работе [7] для восстановления бит ключа 6-раундовоого DES использовалась комбинация машинного обучения и методов дифференциального криптоанализа. В [8] исследовалась возможность применения методов машинного обучения для задачи различения шифртекста криптосистем DES, AES, RC4. В работе [9] – одной из первых, посвященных криптоанализу блочных шифров с помощью нейронных сетей, – исследовалась возможность восстановления ключа для криптосистем, основанных на сети Фейстеля. Рассматривался модельный n-раундовый шифр с длиной блока равной 16 битам. Атака состояла в переборе значений последнего раундового ключа с последующим обучением нейронной сети на множестве пар <открытый текст, шифртекст после n–1 раунда> для аппроксимации (n–1) раунда шифра. Для проверки правильности предположения о последнем раундовом ключе использовались тестовые данные: если предположение о значении последнего раундового ключа неверно, то обученная с таким ключом нейронная сеть на тестовых открытых текстах вместо верных шифртекстов выдаст случайную последовательность бит. Атака на основе предложенной ИНС позволила восстановить ключ 4-раундового шифра. Следует отметить недостаток, присущий многим рассмотренным статьям – недостаточная полнота описания архитектуры используемых нейронных сетей и условий проведения вычислительных экспериментов, что существенно затрудняет возможность воспроизведения приведенных результатов. Например, в работе [10] утверждается об успешном проведении криптоанализа DES и 3DES на основе 2048 и 4096 пар <открытый текст, шифртекст> (полученных на одном ключе) соответственно. Обучившись на указанных данных за сравнительное небольшое время (около 50 минут для DES и 70 минут для 3DES), 4-слойная нейронная сеть, по словам авторов [10], смогла успешно восстанавливать блоки открытого текста по блокам шифртекста без восстановления ключа. Подобные крайне оптимистичные результаты вызвали обоснованные сомнения и в работе [11] воспроизвести результаты из [10] не удалось.

Широко применяется машинное обучение в атаках по сторонним каналам, использующих особенности реализации криптосистем на физическом уровне. Для этих задач используются глубокие нейронные сети [12], метод опорных векторов [13], генеративно-состязательные нейронные сети [14].

В статье [15] предложен поточный шифр на основе нейронной сети, в [16, 17] нейронные сети применяются для анализа генераторов случайных числовых последовательностей. В статье сотрудников Google [18] построена состязательная нейронная сеть, в которой две стороны обучаются шифровать передаваемые друг другу сообщения по прослушиваемому каналу связи.

Методы на основе машинного обучения и нейронных сетей также применяются для анализа трафика. В работе [19] анализировался зашифрованный трафик, генерируемый различными мобильными приложениями: Facebook, Twitter, Dropbox, Gmail и др. с целью идентифицировать действие пользователя в приложении (отправка и открытие сообщения, публикация поста и др.). На рисунке 3 приведен фрагмент статьи [19], иллюстрирующий эффективность предложенного подхода для приложения Facebook.

Apps	Actions	Description	Precision	Recall	F-measure
Facebook	*send message*	send a direct message to a friend	1.00	1.00	1.00
	post user status	post a status on the user's wall	1.00	0.95	0.97
	open user profile	select user profile page from menu	0.96	0.91	0.94
	open message	select a conversation on messages	0.98	1.00	0.99
	status button	select "*write a post*" on user's wall	1.00	1.00	1.00
	post on wall	post a message on a friend's wall	1.00	0.98	0.99
	open facebook	open the Facebook app	1.00	1.00	1.00
	other facebook	other Facebook network traffic	0.99	1.00	0.99
	Average Facebook		0.99	0.98	0.99

Рис. 3. Результаты компьютерных экспериментов [19]

Машинное обучение используется и в стеганографии. Работы [20, 21] посвящены методам стегоанализа на основе нейронных сетей, в статьях [22, 23] ИНС применяются для встраивания секретной информации в контейнер. В [23] построены две глубокие ИНС для встраивания одного цветного изображения – секрета S в другое – контейнер C и последующего извлечения S из C, изображения S и C при этом имеют одинаковый размер. В разработанной системе стеганографической защиты информации встраивание информации может производиться в любые биты, а не только в наименее значимые, как во многих других стеганографических алгоритмах. Вследствие значительного процента информации, встраиваемой в контейнер, метод обладает недостатком – после восстановления в изображениях появляются искажения. На рисунке 4 представлен пример исходных изображений S и C, а также контейнера C' после встраивания в него S и восстановленного секрета S'.

Рис. 4. Пример работы ИНС [23]

В таблице 1 приведены значения ρ(*S, S'*) и ρ(*C, C*'), функция расстояния между изображениями ρ вычислялась как квадратный корень от усредненной по 1000 изображениям разности квадратов каждого пикселя (усредненная сумма квадратов разностей каждого из трех цветовых каналов). Функция ρ, таким образом, характеризует ошибку, вносимую в изображение алгоритмом встраивания, и изменяется от 0 до 255. Параметр β использовался при обучении нейронных сетей, функция ошибки *E=E*(*C*, *C*', *S*, *S*') для которых вычислялась по формуле: $E = \rho(C, C') + \beta\,\rho(S, S')$. Последняя строка в таблице соответствует случаю, когда встраивания в *C* не происходит.

β	ρ(C, C')	ρ(S, S')
0.75	0.75	3.6
1.00	3.0	3.2
1.25	6.4	2.8
0.00	0.1	–

Таблица 1 – Ошибки, вносимые алгоритмом [23]

Таким образом, машинное обучение и ИНС обладают высоким потенциалом для их применения в задачах криптологии и защиты информации – как за счет повышения эффективности существующих методов, так и путем создания на основе ИНС новых методов анализа.

Список литературы

1. Murphy, K. P. Machine Learning: a Probabilistic Perspective. – Cambridge University Press, 2013. – 1104 p.
2. Николенко, С. Глубокое обучение / С.Николенко, А.Кадурин, Е.Архангельская. – СПб.: Питер, 2018. – 480 с.
3. Mitchell, T. M. Machine Learning, 1 edition, New York, NY, USA: McGraw-Hill, Inc., 1997. – 414 p.
4. Rivest, R.L. Cryptography and machine learning / R.L.Rivest // Advances in Cryptology. ASIACRYPT 91. – 1991 – P. 427–439.
5. Danziger, M. Improved cryptanalysis combining differential and artificial neural network schemes / M. Danziger , M. Henriques // 2014 International telecommunications symposium (ITS), IEEE, 2014. – P. 1–5.
6. Gohr,A. Improving attacks on round-reduced speck32/64 using deep learning / A.Gohr // Advances in cryptology, CRYPTO-2019. – 2019. – P. 150—179.
7. Laskari,E. Cryptography and cryptanalysis through computational intelligence / E. Laskari, G. Meletiou, Y. Stamatiou, M. Vrahatis // Computational Intelligence in Information Assurance and Security, Springer. – 2007. – P. 1–49.
8. Chou, J. On the eectiveness of using state-of-the-art machine learning techniques to launch cryptographic distinguishing attacks / J. Chou, S. Lin, C. Cheng // Proceedings of the 5th ACM workshop on Security and articial intelligence, ACM. – 2012. – P 105–110. Albassal, A.M.B. Neural network based cryptanalysis of a Feistel type block cipher / A. M. B. Albassal, A. M. A. Wahdan // International Conference on Electrical, Electronic and Computer Engineering, 2004. ICEEC ‘04. – 2004. – P. 231–237.
9. Alani, M.M. Neuro-cryptanalysis of DES and triple-DES / M.M.Alani // Int. Conf. on Neural information processing. – N.Y.: Springer, 2012. – P.637-646.

10. Xiao, Ya et al. Neural Cryptanalysis: Metrics, Methodology, and Applications in CPS Ciphers. – 2019 IEEE Conference on Dependable and Secure Computing (DSC). –2019. – P. 1-8.
11. Lerman, L. Power analysis attack: an approach based on machine learning / L. Lerman, G. Bontempi, O. Markowitch // International Journal of Applied Cryptography, Vol 3(2). – 2014. – P 97–115.
12. Bartkewitz, T. Template Attacks Based on Probabilistic Multi-class Support Vector Machines / T. Bartkewitz, K. Lemke-Rust // Springer Berlin Heidelberg – 2013. – P.263–276.
13. Wang, P. Enhancing the Performance of Practical Profiling Side-Channel Attacks Using Conditional Generative Adversarial Networks / P. Wang, P. Chen, Z. Luo, G. Dong, M. Zheng, N. Yu, H. Hu. – 2020. – arXiv: 2007.05285.
14. Long, H. Stream Cipher Method Based on Neural Network / H.Long // Proceedings of the 2012 National Conference on Information Technology and Computer Science, CITCS. – 2012. – P. 414–417.
15. Melia-Segui, J. A Practical Implementation Attack on Weak Pseudorandom Number Generator Designs for EPC Gen2 Tags / J. Melia-Segui, J. Garcia-Alfaro, J. Herrera-Joancomarti // Wireless Personal Communications, Vol. 59. – 2011. – P.27-42.
16. Duy, N. Machine learning cryptanalysis of a quantum random number generator / N. Duy, J. Yan, S. Assad, P. Lam, O. Kavehei // IEEE Transactions on Information Forensics and Security, 14(2). – 2018.
17. Abadi, M. Learning to protect communications with adversarial neural cryptography / M. Abadi, D.G.Andersen. – 2016. – arXiv:1610.06918.
18. Conti, M. Analyzing android encrypted network traffic to identify user actions / M. Conti, L. V. Mancini, R. Spolaor, N. V. Verde // IEEE Transactions on Information Forensics and Security, Vol. 11, no. 1. – 2016. – P. 114– 125.
19. Qian, Y. Deep learning for steganalysis via convolutional neural networks / Y. Qian, J. Dong, W. Wang, T. Tan // SPIE/IS&T Electronic Imaging. International Society for Optics and Photonics, 2015.
20. Pibre, L. Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source mismatch / L.Pibre, J. Pasquet, D. Ienco, M. Chaumont // Electronic Imaging, 2016(8). – 2016 – P 1–11.
21. Jarušek, R. Neural network approach to image steganography techniques / R. Jarušek, E. Volna, M. Kotyrba // Mendel, Springer, 2015. – P. 317–327.
22. Baluja, S. Hiding images in plain sight: Deep steganography / S. Baluja // Advances in Neural Information Processing Systems 30. – 2017. – P. 2069– 2079.

УДК 004.056.53

СОВРЕМЕННЫЕ ТРЕНДЫ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ В ОБЛАСТИ ЗАЩИТЫ ИНФОРМАЦИИ

С. Ю. МЕЛЬНИКОВ, В.А. ПЕРЕСЫПКИН
ООО «Лингвистические и информационные технологии», г. Москва, 127018, Российская Федерация
ФГУП «НТЦ «Орион», г. Москва, 127018, Российская Федерация

Введение

Стратегические направления научных исследований в области обеспечения информационной безопасности перечислены в «Доктрине информационной безопасности Россий-