УДК 003.26+004.032.26

ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ В ЗАДАЧАХ ЗАЩИТЫ ИНФОРМАЦИИ

М.В. МАЛЬЦЕВ
Научно-исследовательский институт прикладных проблем математики и информатики г. Минск, Республика Беларусь

Введение. Нейронные сети – активно развивающийся инструмент, доказавший свою эффективность для решения многих практических задач анализа данных. Нейронные сети применяются в задачах распознавания, классификации, анализа данных и прогнозирования, построения систем искусственного интеллекта. В последние годы интерес к нейронным сетям появился и в криптографическом сообществе. Ведутся исследования по применению нейронных сетей для анализа криптографических алгоритмов и протоколов, для разработки и анализа генераторов псевдослучайных чисел, обнаружения вторжений и для других задач. Одной из ключевых особенностей нейронных сетей является отсутствие необходимости четкого знания алгоритма для решения задачи – сеть в процессе обучения самостоятельно вырабатывает наиболее эффективный по заданным критериям алгоритм. Применению нейронных сетей в задачах защите информации посвящена настоящая статья.

1. Искусственные нейронные сети. Идея искусственных нейронных сетей состоит в моделировании работы человеческого мозга, состоящего из множества взаимодействующих между собой нервных клеток – нейронов. У типичного нейрона имеется несколько «входных» отростков – дендритов, через которые он получает сигналы от других нейронов, и один «выходной» отросток – аксон, по которому нейрон передает свой сигнал. Моделируя эту структуру, американский ученый Фрэнк Розенблатт в 1950-х годах предложил конструкцию перцептрона, являющуюся основой нейронных сетей. Схема перцептрона представлена на рисунке 1 [1].

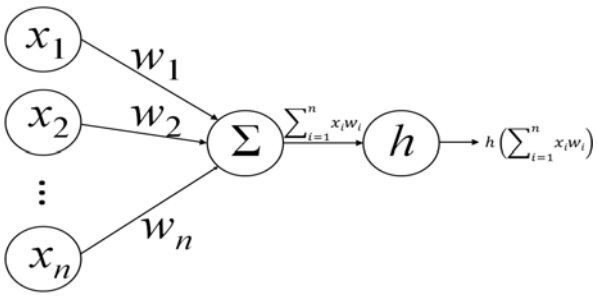

Рис. 1. Схема перцептрона

Здесь $x_1, x_2, ..., x_n$ – входные значения нейрона из некоторого множества (например, из множества действительных чисел □); $w_1, w_2, ..., w_n$ – веса входных значений. Основной задачей при работе с нейронной сетью является ее обучение – определение весов w_i. Функция h называется функцией активации. Распространенным на практике являются следующие типы функций:

– логический сигмоид:

$$h(\mathrm{z}) = \frac{1}{1+e^{-z}};$$

– гиперболический тангенс:

$$h(\mathrm{z}) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}};$$

– функция Хевисайда:

$$h(\mathrm{z}) = \begin{cases} 0, \textit{если } z \le 0, \\ 1, \textit{если } z > 0. \end{cases};$$

– ReLU (rectified linear units):

$$h(\mathrm{z}) = \begin{cases} 0, \textit{если } z \le 0, \\ z, \textit{если } z > 0. \end{cases};$$

Множество соединенных друг с другом перцептронов (как правило, в слои), образуют нейронную сеть. Схема нейронной сети приведена на рисунке 2. В зависимости от строения (архитектуры) выделяют различные типы нейронных сетей: сверточные сети, рекуррентные сети, генеративно-состязательные нейронные сети и другие [1].

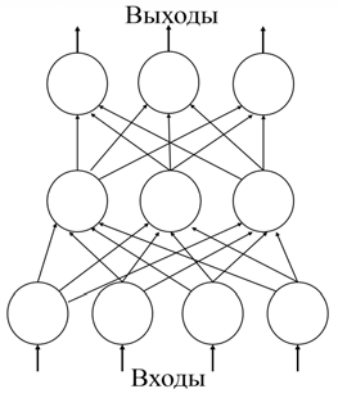

Рис. 2. Схема нейронной сети

2. Применение нейронных сетей в задачах информационной безопасности. Важной практической задачей, решаемой с помощью искусственных нейронных сетей, является классификация. В связи с этим значительное число публикаций посвящено применению нейронных сетей для построения систем идентификации и аутентификации, использующих биометрические характеристики: отпечатки пальцев, изображения лица, радужную оболочку глаза, почерк и др. [2–4]. В [4] сверточные нейронные сети применены для построения системы защиты криптовалютного кошелька: владелец кошелька идентифицируется по изображению его лица; точность распознавания, вычисленная в компьютерных экспериментах, составила 99.3%. Следует отметить, что для разработки методов распознавания пользователей информационных систем по их биометрическим характеристикам требуются достаточно большие базы таких характеристик, получить которые зачастую затруднительно. Решить указанную проблему помогают генеративно-состязательные нейронные сети, которые позволяют создавать необходимые данные. Такие сети, как правило, состоят из двух компонент – генератора и дискриминатора: генератор порождает объекты, принадлежащие различным классам, а дискриминатор пытается отличать объекты из разных классов – в результате такого взаимодействия качество объектов генератора улучшается. В работе [5] генеративно-состязательные нейронные сети применены для генерации отпечатков пальцев.

Применяются нейронные сети и для построения и анализа криптографических алгоритмов и протоколов. В работах [6, 7] нейронные сети используются для конструирования блочных и поточных шифров, в [8, 9] – для функций хэширования. В [9] нейронная сеть с одним скрытым слоем вычисляет 128-битное хэш-значение. Проведено исследование криптографических свойств построенной функции: стойкость к атаке нахождения прообраза, к атаке «дней рождения». Наличие лавинного эффекта проиллюстрировано на рисунке 3: по горизонтальной оси откладывался номер бита, изменяемого в сообщении, по вертикальной оси – расстояние Хэмминга между хэш-значениями исходного сообщения и измененного. Расстояние Хэмминга колеблется около 50, что говорит о наличии лавинного эффекта: изменение одного бита влечет изменение около половины бит хэш-значения. Предложенная в [9] хэш-функция допускает распараллеливание, что позволило добиться превосходства в скорости над хэш-функциями MD5 и SHA-1.

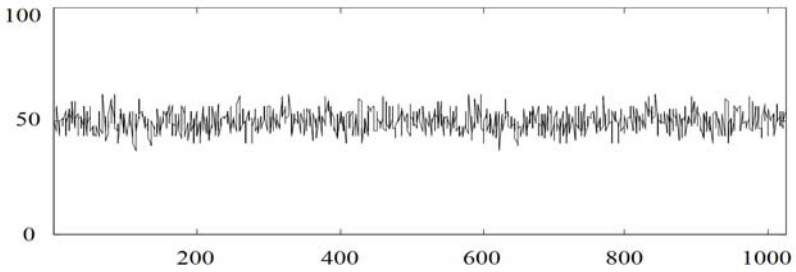

Рис. 3. Зависимость хэш-значения от изменения бит сообщения

Ряд публикаций посвящен применению нейронных сетей для криптоанализа систем шифрования [10–12]. Статья [11] посвящена атакам по сторонним каналам, использующим особенности практической реализации криптографических алгоритмов, которые могут приводить к уязвимостям. Анализируются показатели физических процессов, происходящих в конкретном устройстве, на котором реализован криптографический алгоритм: время выполнения алгоритма, потребляемая мощность, напряжение и сила тока в узлах устройства и другие характеристики. На основании этих данных становится возможным за сравнительное небольшое время извлекать информацию о секретных параметрах криптосистемы. В [11] применение сверточных нейронных сетей позволило повысить эффективность подобных атак.

В работе [12] нейронные сети применялись для распознавания шифртекстов различных криптосистем. Для этого использовались каскадные корреляционные нейронные сети (cascade correlation neural networks) и сети с обратным распространением ошибки (back propagation networks). Анализировались выходные последовательности блочного шифра RC6e (enhanced

RC6) и поточного шифра SEAL. В ходе вычислительных экспериментов использовался шифртекст длиной 1,92 миллиона символов, из которых 1,66 миллиона использовались для обучения нейронной сети, остальные 260 тысяч – для проверки качества распознавания. Нейронные сети позволили достичь высокой точности классификации: для сети с обратным распространением ошибок она составила 85%, для каскадной корреляционной сети – 92%.

В завершение приведем работу сотрудников компании Google [13], в которой исследовалась возможность построения нейронными сетями криптографических алгоритмов, обеспечивающих конфиденциальную передачу информации. Для этой цели использовались уже упоминаемые в настоящей статье генеративно-состязательные нейронные сети. Моделировались три сети: Алиса, Боб и Ева. Алиса и Боб, имеющие общий секретный ключ, обмениваются сообщениями по открытому каналу связи, прослушиваемому Евой. Цель Алисы и Боба состоит в обеспечении конфиденциальности их переписки, цель Евы – нарушить эту конфиденциальность. Отметим, что никакого априорного знания алгоритмов шифрования во взаимодействующие нейронные сети не закладывалось. В результате Алиса и Боб обучаются симметричному шифрованию, «изобретая» некоторое подобие одноразового блокнота. Результаты вычислительных экспериментов представлены на рисунке 4. Передаваемые секретные сообщения представляли собой случайные 16-битовые последовательности. Задача Боба и Евы состояла в восстановлении переданного Алисой сообщения. На рисунке 4 представлены графики зависимости ошибок Боба и Евы от числа итераций обучения. Видно, что со временем Боб обучается безошибочно расшифровывать сообщение, а Ева ошибается в восьми битах, что соответствует случайному угадыванию открытого текста.

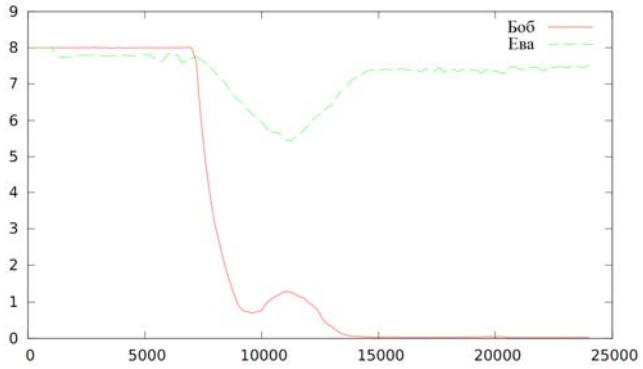

Рис. 4. Зависимость ошибок от числа итераций обучения

Список литературы

1. Николенко, С. Глубокое обучение / С. Николенко, А. Кадурин, Е. Архангельская. – СПб. : Питер, 2018. – 480 с.

2. Rathgeb, C. A survey on biometric cryptosystems and cancelable biometrics / C. Rathgeb, A. Uhl // EURASIP Journal on Info. Security. – 2011. – № 3.

3. Tarek, M. Robust cancellable biometrics scheme based on neural networks / M. Tarek, O. Ouda, T.Hamza // IET Biometrics. – 2016. – № 5 – P.220–228.

4. Albakri, A. Convolutional neural network biometric cryptosystem for the protection of the blockchain's private key / A. Albakri, C. Mokbel // Procedia Computer Science. – 2019. – Vol. 160. – P. 235–240.

5. Riazi, M. Automatic Synthetic Fingerprint Generation / M. Riazi, S. Chavoshian, F. Koushanfar . – 2020.

6. 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.

7. Lian, S. A block cipher based on chaotic neural networks / S. Lian // Neurocomputing. – 2009. – Vol. 72. – P. 1296–1301.

8. Turcanik, M. Hash function generation by neural network / M. Turcanik, M. Javurek // New Trends in Signal Processing (NTSP). – 2016. – P. 1–5.

9. Lian, S. One-way Hash Function Based on Neural Network / S.Lian, J.Sun, Z. Wang. – 2007.

10. Xiao, Y. Neural Cryptanalysis: Metrics, Methodology, and Applications in CPS Ciphers / Y. Xiao, Q. Hao, D. Yao // 2019 IEEE Conference on Dependable and Secure Computing (DSC). – 2019. – P. 1–8.

11. Cagli, E. Convolutional neural networks with data augmentation against jitter-based countermeasures / E. Cagli, C. Dumas, E. Prouff // International Conference on Cryptographic Hardware and Embedded Systems. – 2017. – P. 45–68.

12. Chandra, B. Applications of cascade correlation neural networks for cipher system identification / B. Chandra, P. Varghese // World Academy of Science, Engineering and Technology, Vol. 26. P. 312–314, 2007.

13. Abadi, M. Learning to protect communications with adversarial neural cryptography / M. Abadi, D. G. Andersen. – 2016.