НАУЧНОЕ ИЗДАНИЕ МГТУ ИМ. Н. Э. БАУМАНА
НАУКА и ОБРАЗОВАНИЕ
Эл № ФС77 - 48211. Государственная регистрация №0421200025. ISSN 1994-0408
электронный научно-технический журнал

Нейросетевая защита ресурсов автоматизированных систем от несанкционированного доступа

 # 05, май 2013
DOI: 10.7463/0513.0566210
Булдакова Т. И.
УДК 004.056.53

Россия, МГТУ им. Н.Э. Баумана
buldakovati@gmail.com

Введение

Важной и ответственной задачей при создании и эксплуатации автоматизированных систем является обеспечение их информационной безопасности (ИБ). Несанкционированные преднамеренные или случайные воздействия на информационные ресурсы могут привести к существенным нарушениям в работе разнообразных транспортных, энергетических, промышленных, финансовых и других автоматизированных систем [1]. Угрозы безопасности и атаки могут быть направлены на нарушение конфиденциальности, целостности и доступности данных, а также могут привести к нарушению работоспособности самих систем.

В настоящее время наиболее перспективным направлением разработок в сфере ИБ является использование возможностей искусственного интеллекта для предотвращения угроз и предсказания событий, которые могут привести к сбою систем. Благодаря гибкости и возможности обработки больших массивов данных, интеллектуальные технологии являются хорошим инструментом для обеспечения информационной безопасности ресурсов.

1 Постановка задачи

Для ограничения доступа к ресурсам активно применяют разнообразные системы парольной защиты, замково-ключевые устройства (смарт-карты, магнитные карты), биометрические системы (отпечатки пальцев, снимки сетчатки глаза, голосовая идентификация), криптографические системы защиты (электронная подпись) [2]. Однако, этого зачастую оказывается недостаточно.

По мере развития информационных технологий и средств похищения конфиденциальной информации у злоумышленников появилась возможность применить

новые способы для получения доступа к ресурсам, используя чужие данные при идентификации. Чаще всего используются специально обученные «программы-роботы», которые проводят массовые регистрации на открытых ресурсах, производя огромную бесполезную нагрузку на вычислительные мощности и угрожая стабильности систем, или, обладая похищенной информацией, производят авторизацию на закрытые ресурсы с целью похищения полезной информации.

Средствами защиты от данных типов угроз являются встраиваемые модули, проверяющие пользователя и выявляющие подобные автоматизированные программы. Одним из средств проверки является генерация изображений с информацией, которую предлагается ввести пользователю для доказательства того, что он является человеком. В связи с тем, что изображения повторяют друг друга по типу искажений, злоумышленники стали активно использовать искусственные нейронные сети (ИНС) для распознавания данных путем поэтапного удаления шумов с изображений. Поскольку регистрация на всех ресурсах выполняется одинакового, с незаметными отступлениями, то «роботу» с помощью ИНС не составляет труда пройти регистрацию и получить легальный доступ в систему.

Поэтому необходимо расширить возможности встраиваемого модуля защиты за счет включения в него алгоритма авторизации пользователя по сгенерированному изображению, которое сможет распознать человек и не сможет распознать «программа-робот». В качестве своеобразного фильтра, позволяющего отбрасывать образы с высокой вероятностью распознавания, предлагается использовать нейронную сеть.

2 Решение

В зависимости от операций над образом он может быть эталонным, искаженным или распознанным. Введем обозначения: P_{init} – эталонный образ; P_{dist} – искаженный образ; P_{recogn} – распознанный образ. Искаженный и эталонный образы связаны соотношением $P_{dist} = F(P_{init})$, где F – некоторая композиция функций искажения изображения f_i, $i = [1, m]$. Здесь m – количество базовых искажений.

Генерация эталонного образа состоит в выборе символов, а генерация искаженного образа – в случайном выборе одного или нескольких базовых искажений и их применении к эталонному образу.

В соответствии с предлагаемым алгоритмом, если при распознавании изображения нейронная сеть определит эталонный образ, т.е. $P_{\text{recogn}} = P_{init}$, то данное изображение отбрасывается. В случае $P_{\text{recogn}} \neq P_{init}$ искаженный образ передается для распознавания пользователю (рис. 1).

Рис. 1. Основные этапы алгоритма распознавания образов

Для генерации образов используются буквы латинского алфавита без учета регистра и цифры (0 - 9). В результате на выходе имеется растровое изображение из 4 - 6 символов, выбранных произвольно и располагающихся в центре на белом фоне. Размер изображения 100 × 100 пикселей. Таким образом, изображение представляет собой набор точек $P(x, y)$ с координатами x и y, где $x = [1, 100]$, $y = [1, 100]$.

Полученное изображение считается эталонным, и на него накладываются необходимые искажения, которые требуются для уменьшения вероятности распознавания искаженного образа интеллектуальной «программой-роботом».

Базовые искажения должны быть достаточно простыми, чтобы их можно было комбинировать. Поэтому в качестве основных выберем следующие искажения:

- волновое искажение $Dist_{wave}(P(x, y))$;
- смещение символов $Dist_{displacement}(P(x, y))$;
- поворот символов $Dist_{rotate}(P(x, y))$;
- шумы в виде точек $Dist_{point\_noise}(P(x, y))$;
- шумы в виде линий $Dist_{line\_noise}(P(x, y))$;
- склейка символов $Dist_{imposition}(P(x, y))$.

Примеры этих видов искажений приведены на рис. 2.

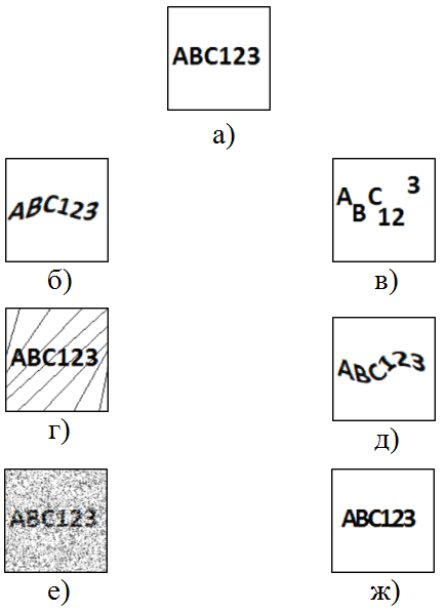


Рис. 2. Виды искажений: а) исходное изображение; б) волновое искажение; в) смещение символов; г) шум линиями; д) поворот; е) шум точками; ж) склейка символов

Кроме того, необходимо выбрать тип используемой нейронной сети. Поскольку для ограничения доступа к ресурсам требуется распознавать искаженные образы, то целесообразно использовать динамические сети, например, сеть Хопфилда [3].

Нейронная сеть Хопфилда – это полносвязная сеть с симметричной матрицей связей. В процессе работы динамика таких сетей сходится (конвергирует) к одному из положений равновесия. Эти положения равновесия являются локальными минимумами функционала, который называется «энергия сети». Такая сеть может быть использована как ассоциативная память, как фильтр, а также для решения некоторых задач оптимизации.

В отличие от многих нейронных сетей, работающих до получения ответа через определённое количество тактов, сети Хопфилда работают до достижения равновесия, когда следующее состояние сети в точности равно предыдущему: начальное состояние является входным образом, а при равновесии получают выходной образ.

Задача, решаемая данной сетью в качестве ассоциативной памяти, как правило, формулируется следующим образом. Известен некоторый набор двоичных сигналов (изображений, звуковых оцифровок, прочих данных, описывающих некие объекты или характеристики процессов), которые считаются эталонными. Сеть должна уметь из произвольного неидеального сигнала, поданного на ее вход, выделить («вспомнить» по частичной информации) соответствующий образ (если такой есть) или «дать заключение» о том, что входные данные не соответствуют ни одному из образов [4].

Структурная схема сети Хопфилда приведена на рис. 3. Она состоит из единственного слоя нейронов, число которых является одновременно числом входов и выходов сети. Каждый нейрон связан синапсами со всеми остальными нейронами, а также имеет один входной синапс, через который осуществляется ввод сигнала. Выходные сигналы образуются на аксонах.

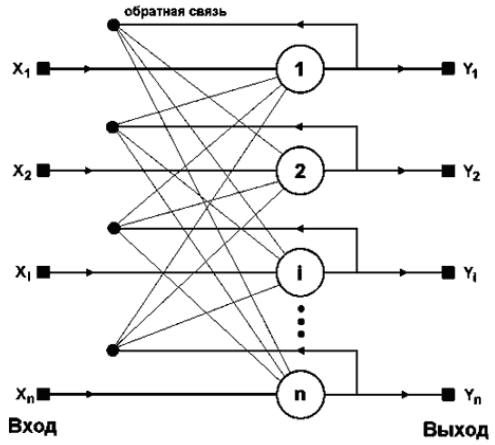


Рис. 3. Структурная схема сети Хопфилда

В этой модели состояние системы описывается n-мерным вектором $V = (V_1,...V_n)$, где V_i - одна из вершин единичного гиперкуба в n-мерном пространстве меняется во времени, 0 или 1 описывает состояние i - го нейрона.

Элементы сети функционируют в асинхронном режиме, то есть каждый нейрон в случайные моменты времени с некоторой частотой определяет свое состояние. Это позволяет описать поведение сети как релаксационный процесс, при котором минимизируется энергетическая функция E модели

$$E = -\frac{1}{2}\sum_{i\neq j}\sum W_{ij}V_iV_j + \sum_i Q_iV_i,$$

где W_{ij} - матрица связей; V и Q - состояние и порог нейрона.

Изменение энергии сети E при изменении состояния нейрона равно

$$\Delta E = -\Delta V_i \sum_i W_{ij}V_j.$$

Эволюция системы из произвольного начального состояния может закончиться только в одной из стационарных точек, соответствующих локальному минимуму энергии E.

Алгоритм функционирования сети Хопфилда можно представить состоящим из двух шагов.

На первом шаге формируется синаптическая карта W сети путем ее обучения по серии входных образов $A^k = [a_1^k, \ldots, a_n^k]$, $k = 1, \ldots, L$, где L – количество входных образов. Для бинарных векторов элементы матрицы равны

$$W_{ij} = \sum_k \left(2V_i^k - 1\right)\left(2V_j^k - 1\right),$$

а для биполярных векторов –

$$W_{ij} = \sum_k V_i^k V_j^k.$$

Таким образом, в каждом элементе синаптической карты содержится информация обо всех запоминаемых образах. В модели Хопфилда дополнительно обнуляется диагональ, т.е. $W_{ii} = 0$, матрица W является симметричной.

На втором шаге происходит собственно распознавание образов. На вход сети подается некоторый образ $C = [c_1, \ldots, c_n]$, т.е. сеть приводится в состояние $V_j = c_j$, $j = 1, \ldots, n$. Далее осуществляется итерационное вычисление выходных сигналов сети до тех пор, пока сеть не достигнет установившегося состояния:

$$V_j(t+1) = f\left(\sum_{i=1}^{n} V_i(t) W_{ij}\right),\ j = 1,\ldots, n.$$

Таким образом, если сеть Хопфилда включить в модуль защиты от несанкционированного доступа, то ее можно использовать в качестве фильтра, отсеивающего распознанные образы, что обеспечит значение показателей эффективности защиты на необходимом уровне.

3 Полученные результаты

Для проверки эффективности алгоритма он был реализован в виде программного продукта, который выполняет следующие функции:

- формирование изображения;
- нанесение шумов на изображение;
- обучение нейронной сети Хопфилда;
- восстановление эталонного образа из искаженного.

Поскольку на вход сети подается растровое изображение размером 100×100 пикселей, а емкость сети Хопфилда составляет 15% от количества используемых нейронов, то максимальное количество образов, которое она может запомнить, равно 1500. Однако при образах, которые можно считать похожими, сеть может проводить распознавание ошибочно. Кроме того, при использовании композиции искажений нейронной сети будет сложнее распознать образ, но для человека это не должно составить труда (рис. 4).

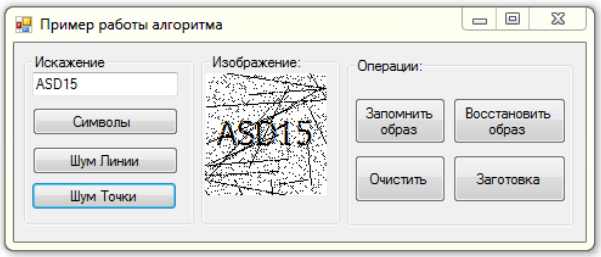
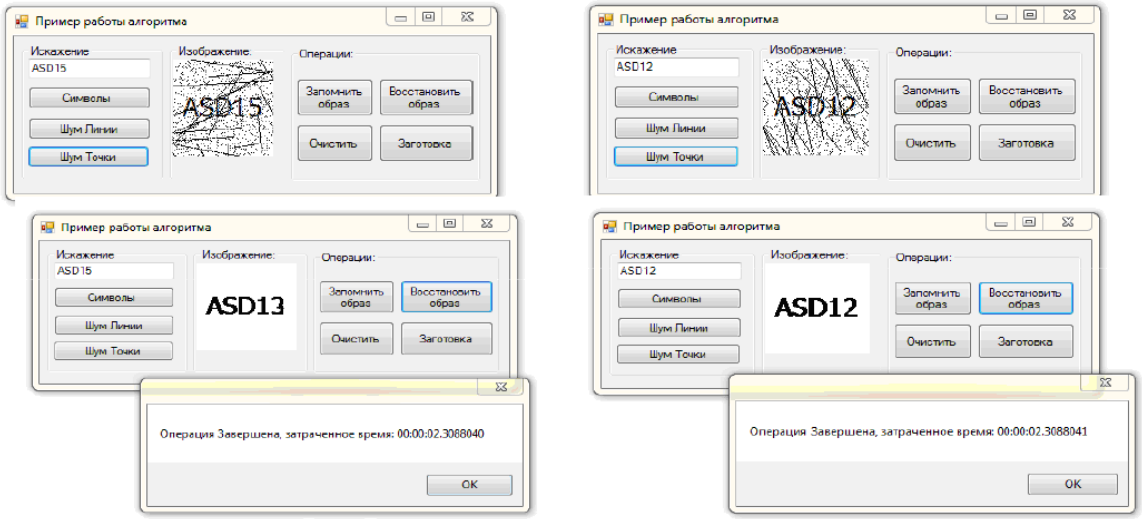


Рис. 4. Использование композиции искажений

На рис. 5 приведены примеры распознавания сетью Хопфилда искаженных образов. На левом рисунке сеть не смогла правильно распознать изображение, поэтому его можно передавать для авторизации пользователя. На правом рисунке сеть восстановила искаженное изображение, следовательно, его нужно отфильтровать, поскольку нейронная сеть злоумышленника также может это сделать.



а) б)

Рис. 5. Пример распознавания изображения: а) сеть не распознала образ; б) правильное распознавание образа

При верификации алгоритма на тестовых примерах была собрана статистическая информация, включающая в себя количество распознанных образов. Отметим, что при формировании искаженных образов применялись одно или несколько искажений. Полученная информация приведена в табл. 1.

Таблица 1.

Результаты тестирования алгоритма

Кол-во образов поданных на вход ИНС	Количество верно распознанных образов	Количество неверно распознанных образов	Вероятность распознавания образов
5	5	0	1
10	9	1	0,9
20	17	3	0,85
30	23	7	0,77

Результаты исследования показали, что при обучении искусственной нейронной сети эталонным образам, в которых присутствуют соответствия с другими образами (например, одинаковые символы, расположенные в тех же позициях), происходит неверное распознавание. Число неверных распознаваний увеличивается с числом запомненных нейронной сетью эталонных образов. Кроме того, на вероятность распознавания образов влияют искажения: при использовании более одного искажения вероятность правильного распознавания уменьшается.

Заключение

Сеть Хопфилда позволяет просто и эффективно решить задачу распознавания образов по неполной и искаженной информации. Невысокая емкость сетей (число запоминаемых образов) объясняется тем, что сети не просто запоминают образы, а позволяют проводить их обобщение. Вместе с тем, легкость построения программных и аппаратных нейросетевых моделей делают эти сети привлекательными для многих применений, в том числе для обеспечения информационной безопасности автоматизированных систем.

Список литературы

1. Норенков И.П. Автоматизированные информационные системы: учеб. пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011. 342 с.

2. Мельников В.В. Безопасность информации в автоматизированных системах. М.: Финансы и статистика, 2003. 368 с.

3. Чулюков В.А., Астахова И.Ф., Потапов А.С., Каширина И.Л., Миловская Л.С., Богданова М.В., Просветова Ю.В. Системы искусственного интеллекта. М.: БИНОМ. Лаборатория знаний, 2008. 292 с.

4. Хайкин С. Нейронные сети: полный курс : пер. с англ. М.: Издательский дом «Вильямс», 2006. 1104 с.

SCIENTIFIC PERIODICAL OF THE BAUMAN MSTU
SCIENCE and EDUCATION
EL № FS77 - 48211. №0421200025. ISSN 1994-0408
electronic scientific and technical journal

Neural network protection of automated systems' resources from unauthorized access

 # 05, May 2013
DOI: 10.7463/0513.0566210
Buldakova T.I.

Bauman Moscow State Technical University, 105005, Moscow, Russian Federation
buldakovati@gmail.com

The problem of information security in automated systems is considered in this article. Various approaches to restriction of access to information resources were analyzed. An authorization algorithm was developed; the algorithm uses images which a human will be able to recognize but which an intellectual "program robot" will not be able to recognize. Basic types of distortions of reference images were chosen. It is proposed to apply a dynamic neural network as a peculiar filter allowing one to reject images with high probability of recognition. Hopfield's recurrent network was used in the implementation of the algorithm.

Publications with keywords: pattern recognition, neural networks, information technology security, resources, automated systems
Publications with words: pattern recognition, neural networks, information technology security, resources, automated systems

References

1. Norenkov I.P. *Avtomatizirovannye informatsionnye sistemy* [The automated information systems]. Moscow, Bauman MSTU Publ., 2011. 342 p.

2. Mel'nikov V.V. *Bezopasnost' informatsii v avtomatizirovannykh sistemakh* [Security of information in the automated systems]. Moscow, Finansy i statistika, 2003. 368 p.

3. Chuliukov V.A., Astakhova I.F., Potapov A.S., Kashirina I.L., Milovskaia L.S., Bogdanova M.V., Prosvetova Iu.V. *Sistemy iskusstvennogo intellekta* [Systems of artificial intelligence]. Moscow, BINOM. Laboratoriia znanii, 2008. 292 p.

4. Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd ed. Prentice Hall, 1999. 823 p. (Russ. ed.: Khaikin S. *Neironnye seti: polnyi kurs*. Moscow, Izdatel'skii dom «Vil'iams», 2006. 1104 p.).