А. Н. Атаманов

МОДУЛЬ НЕЧЕТКОГО ВЫВОДА НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ДИНАМИЧЕСКОГО ИТЕРАТИВНОГО АНАЛИЗА РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В условиях увеличивающейся сложности автоматизированных систем вопросы обеспечения информационной безопасности приобретают все большее значение для государства и бизнеса. Особое внимание начинает уделяться анализу рисков информационной безопасности как необходимой составляющей комплексного подхода к обеспечению ИБ. В связи с выходом закона «О персональных данных» и принятием соответствующих нормативных и методических документов данная процедура стала обязательной для всех операторов персональных данных, т. е. практически для всех организаций и ведомств, использующих информационные технологии. Этап анализа рисков также является обязательным в соответствии со стандартом ISO 27001 (ГОСТ Р ИСО/МЭК 27001), NIST 800 и рядом других международных, национальных и отраслевых стандартов.

Как следствие большого количества стандартов и подходов, основные понятия в этой области имеют множество определений. Наиболее подходящим для большинства практических применений определением риска информационной безопасности является данное в стандарте ISO 27005. Согласно этому стандарту, «риск информационной безопасности — это потенциальная возможность использования уязвимостей актива или группы активов конкретной угрозой для причинения ущерба организации». Из этого определения следует, что риск — это комплексная величина, определяемая как функция (или функционал) ряда других. Сложности проведения анализа рисков напрямую вытекают из трудностей и ошибок при анализе составляющих риска. Помимо организационных вопросов, можно выделить следующие основные сложности:

— заведомая неполнота информации о составляющих риска и их неоднозначные свойства;

— сложность создания модели информационной системы;

— длительность процесса и быстрая потеря актуальности результатов оценки;

— сложность агрегации данных из различных источников, в том числе статистик и экспертных оценок;

— необходимость привлечения отдельных специалистов по анализу рисков.

В связи с этим особую актуальность представляют активно развивающиеся методы непрерывного аудита и анализа рисков информационной безопасности. Вместе с современными моделями управления информационными системами, системами менеджмента информационной безопасности, мониторинга и анализа защищенности данные методологии позволяют наиболее быстро и эффективно строить и развивать систему защиты информации организации. Система непрерывного динамического аудита и анализа рисков позволяет специалистам проводить итеративную оценку рисков с учетом имеющихся данных по бизнес-ландшафту, актуальной информации по используемым или предполагаемым к внедрению технологиям, имеющимся или возможным уязвимостям и их вероятностям.

Особую роль в непрерывном анализе рисков при этом должна играть функция прогнозирования рисков, связанных с планируемыми к внедрению технологиями. Путем автоматизации процесса учета угроз, связанных с появлением новых уязвимостей в типовом ПО, формализации изменений в бизнес-ландшафте и информационной системе, агрегации данных из различных источников можно создать среду, позволяющую специалисту формировать отчеты о состоянии защищенности той или иной информационной системы, основываясь на серии последовательных отчетов, составленных за короткий промежуток времени. Обработка этих данных с использованием методов статистического прогнозирования позволит определить оптимальный набор контрмер с учетом «будущих рисков» и тем самым повысить эффективность внедрения превентивных контрмер и существенно снизить время реакции системы на появление новых уязвимостей [1].

Для создания такой среды необходимо построение модели информационной системы, что само по себе является сложной задачей, требующей, как правило, существенных упрощений. Однако применение теории нечетких множеств для решения этой задачи не предполагает знание модели информационной системы. В этом случае следует сформулировать только правила поведения в форме нечетких условных суждений типа **ЕСЛИ... ТО**. При решении задачи анализа рисков информацию, необходимую для осуществления процесса, можно разделить на две части: численную (количественную) и лингвистическую (качественную), поступающую от экспертов. Значительное количество нечетких систем использует только второй вид знаний, обычно представляемых в виде базы нечетких правил. В случае, когда возникает необходимость спроектировать систему, но в наличии имеются только численные или смешанные данные, задача существенно усложняется. Одним из путей решения считаются так называемые нейро-нечеткие (fuzzy-neural) системы [2].

Рассмотрим созданную модель нейро-нечеткой системы анализа рисков. Входом системы будут данные датчиков (например, системы обнаружения вторжений, антивирусы, межсетевые экраны) о потенциально опасной активности, общем уровне сетевой активности и нагрузки на тот или иной участок автоматизированной системы и т. д., а также экспертные оценки количественных показателей функционирования системы информационной безопасности. Выходом будет являться количественная оценка рисков информационной безопасности. Например, она может быть представлена в виде среднегодовых потерь организации в результате инцидентов безопасности (Annual Loss Expectancy, ALE).

Из определения риска следует, что величина риска R есть функция от ресурса (актива) A, угрозы T и уязвимости V [4].

$R = f(A, T, V)$.

Вопросу использования нечетких множеств для анализа рисков посвящены работы [4–7]. Рассмотрим отдельно модуль нечеткого вывода на основе нейронных сетей. На рисунке 1 изображена а качестве примера схема этого модуля для двух входных переменных.

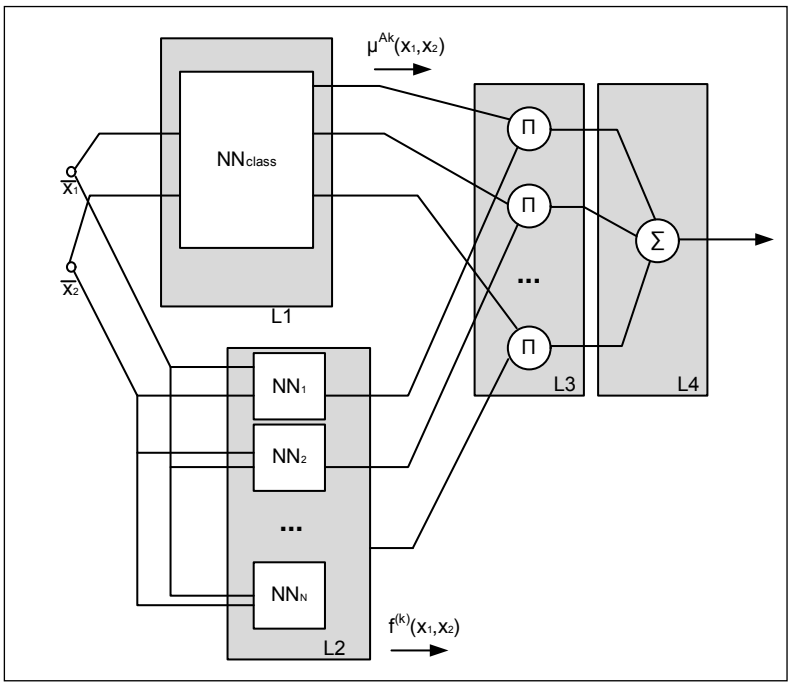

Рис. 1. Модуль нечеткого вывода

Рассмотрим блок, применяемый для решения задачи кластеризации (clustering), обозначенный на схеме NN_{class} (слой L_1). Его задача заключается в разделении множества входных данных на N

нечетких классов, каждый из которых выступает в роли условия для одного из нечетких правил. Этот блок реализуется с помощью нейронной сети, имеющей n входов, N выходов и K слоев. Для обучения сети на вход подаются наборы входных данных и одновременно информация о том, к какому классу принадлежит каждое поданное на вход значение. На втором этапе проверяется то, насколько корректно обучена сеть.

Блоки, обозначенные NN_1 ... NN_N, реализуют заключения соответствующих нечетких правил. Для этого также используются нейронные сети. Каждая сеть имеет n входов и один выход. Эти сети обучаются аналогично, но только после сети NN_{class}, так как имеет существенное значение то, какая сеть должна обучаться по конкретной реализации входного вектора.

Для обучения представленной структуры входные данные должны быть разделены на репрезентативные классы, на базе которых будут формироваться правила. Непосредственное обучение систем будем производить с помощью рекуррентного метода наименьших квадратов. Структура нейрона при этом будет иметь вид, изображенный на рисунке 2.

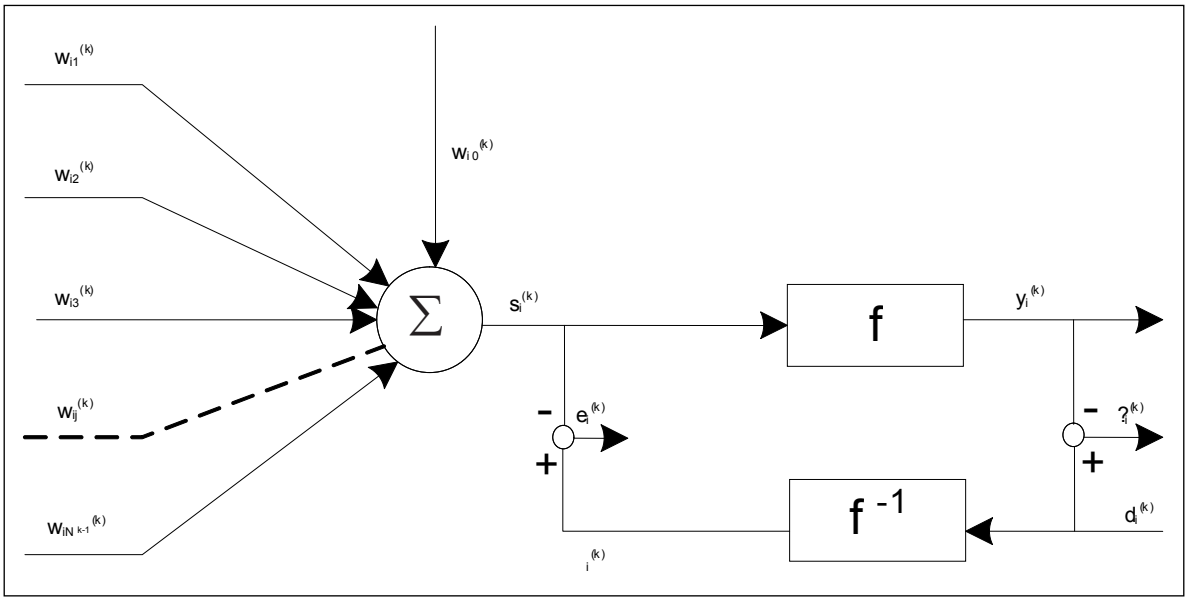

Рис. 2. Структура нейрона

В качестве меры погрешности выходной оценки рисков по отношению к результату анализа инцидентов за прошедшее время используется выражение

$$Q(n)=\sum_{t=1}^{n}\lambda^{n-t}\sum_{j=1}^{N_L}\varepsilon_j^{(L)2}(t)=\sum_{t=1}^{n}\lambda^{n-t}\sum_{j=1}^{N_L}[d_j^{(L)}(t)-f(x^{(L)T}(t)w_j^{(L)}(n))]^2, \quad (1)$$

где λ — коэффициент забывания (forgetting factor).

Выходы блока NN_{class} модуля нечеткого вывода соединены с выходами блоков NN_k посредством мультипликации. Сумма их выходов определяет количественное значение риска, определяемого по формуле

$$\bar{y} = \sum_{k=1}^{N} [\left[\mu_{Ak}\right] (\bar{x}_1, ..., \bar{x}_n)\cdot f^{(k)}(\bar{x}_1, ..., \bar{x}_n)\right], \quad (2)$$

где $f^{(k)}(\bar{x}_1, ..., \bar{x}_n)$ — это результаты функционирования сети NN_k, а μ_{Ak} $(\bar{x}_1, ..., \bar{x}_n)$ — выходы сети NN_{class}, интерпретируемые как принадлежность входных данных соответствующему нечеткому множеству (классу).

В результате получаем выражение, соответствующее методу дефуззификации, путем нахождения центра тяжести (center average defuzzification):

$$\bar{y} = \frac{\sum_{k=1}^{N} \left[\left[\mu_{A^k}\right](\bar{x}_1, ..., \bar{x}_n) \cdot f^{(k)}(\bar{x}_1, ..., \bar{x}_n)\right]}{\sum_{k=1}^{N} \mu_{A^k}(\bar{x}_1, ..., \bar{x}_n)}. \quad (3)$$

Таким образом, на выходе системы получаем оценку риска информационной безопасности на основе имеющихся входных данных. При этом модуль обладает способностью к обучению, т. е. уточнению оценки риска по результатам сравнения с результатами анализа инцидентов безопасности за прошедший период. Помимо этого, данный модуль оказывается отлично приспособленным для проведения динамического анализа рисков и итеративного выполнения этапа анализа. Применение модуля нечеткого вывода на основе нейронных сетей позволяет решить проблему заведомой неполноты информации о составляющих риска и их неоднозначных свойств. При применении разработанной модели отсутствует необходимость в создании детальной модели информационной системы для проведения анализа рисков. Система позволяет агрегировать данные из различных источников и хорошо приспособлена для итеративного динамического анализа рисков.

СПИСОК ЛИТЕРАТУРЫ:

1. *Атаманов А. Н., Минаева Е. В.* Мониторинг информационных рисков как средство повышения защищенности информационных систем // Российская научная конференция «Методы и средства обеспечения информационной безопасности». СПб., 2008. С. 97.

2. *Рутковская Д., Пилиньский М., Рутковский Л.* Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И. Д. Рудинского. М.: Горячая линия — Телеком, 2008. — 452 с.: илл.

3. *Астахов А. М.* Искусство управления информационными рисками. М.: ДМК Пресс, 2010. — 312 с.: илл.

4. *Сидоров А. О.* Метод и модель структурированной оценки риска при анализе информационной безопасности. Дисс. … канд. техн. наук. М.: СПбГУ ИТМО, 2008. — 134 с.

5. *Sodiya A. S., Onashoga S. A., Oladunjoye B. A.* Threat Modeling Using Fuzzy Logic Paradigm // Issues in Informing Science and Information Technology. 2007. Vol. 4. — 51 p.

6. *McGill W. L., Ayyub B. M.* Multicriteria Security System Performance Assessment Using Fuzzy Logic / Journal of Defense Modeling and Simulation. 2007. Vol. 4. No. 4.

7. *Kyoomarsi F., Khosravyan Dehkordy P., Hosein Peiravy M., Heidary E.* Using UML and Fuzzy Logic in optimizing risk management modeling // IADIS International Conference Informatics. 2008. — 9 p.