Инженерный вестник Дона, №4(2017)
ivdon.ru/ru/magazine/archive/n4y2017/4432

Подход к оценке DLP-систем с использованием средств нечеткой логики

А.Р. Айдинян, О.Л. Цветкова

Донской государственный технический университет, Ростов-на-Дону

Аннотация: В настоящее время на рынке программного обеспечения представлено множество многообразных DLP-систем, предназначенных для защиты конфиденциальной информации от утечек. Поскольку подобные системы имеют относительно высокую стоимость, целесообразным является выбор оптимального варианта до приобретения и внедрения, поскольку неправильный выбор приведет к лишним расходам и неэффективному контролю информации. В работе приведено описание базы правил системы нечеткого вывода, предназначенной для оценки DLP-системы. Практическое использование предложенной нечеткой системы позволить выполнить аргументированный выбор такой DLP-системы, которая наиболее полно удовлетворяет запросам конкретной организации и обеспечивает требуемый уровень защиты информации от утечек. При этом оценивание выполняется по нескольким критериям, значения которых, полученные на основе экспертных оценок, являются входными переменными системы нечеткого вывода. На выходе системы нечеткого вывода формируется комплексная оценка DLP-системы, которая отражает степень ее пригодности для внедрения на конкретном предприятии.
Ключевые слова: информационная безопасность, защита информации, защита конфиденциальной информации, DLP-система, искусственный интеллект, нечеткая логика, система нечеткого вывода.

Введение и постановка задачи

В настоящее время, в связи с распространением использования сети Интернет, компьютерных сетей и различных носителей информации, компании терпят убытки, возникающие в результате утечки конфиденциальной информации. Это делает задачу защиты информации компании от утечек весьма актуальной [1]. Риски, связанные с разглашением или утерей конфиденциальной информации, заставляют компании внедрять DLP-системы —системы защиты конфиденциальной информации от утечек.

В связи с достаточно большим многообразием DLP-систем, представленных на рынке программного обеспечения, и их высокой стоимостью, целесообразным является выбор наиболее предпочтительной для использования в компании до приобретения и внедрения. Для выбора DLP-системы предлагается использовать средства искусственного интеллекта. В настоящее время искусственный интеллект с успехом приме-

няется при решении различных задач и реализуется в виде искусственных нейронных сетей [2-4], нечеткой логики [5-7], генетических алгоритмов [8-10], и т.д. В данной работе рассмотрена задача формирования базы продукционных правил для построения системы нечеткого вывода, предназначенной для комплексной оценки DLP-систем. В работах [11, 12] используются различные подходы для получения комплексной оценки в различных областях применения.

Разработка критериев оценки эффективности DLP-систем

Для оценки DLP-систем предлагается использовать ряд критериев [13], каждый из которых описывается с помощью лингвистической переменной и является входной переменной системы нечеткого вывода:

1. Технологии распознавания конфиденциальной информации (код лингвистической переменной *A1*). На сегодня существует несколько базовых технологий, каждая из которых подходит для анализа информации различных видов в том числе и для поиска информации в зашифрованных данных [14].

2. Полнота контролируемых каналов (*A2*). Каждый канал передачи информации — это потенциальный канал утечек,поэтому важно блокировать все неиспользуемые для работы каналы, а оставшиеся — контролировать.

3. Удобство управления (*A3*). Для сохранения качества распознавания конфиденциальной информации на должном уровне необходимо осуществлять регулярное обслуживание, аудит и корректировки настроек. С этой целью DLP-система должна гарантировать простоту установки, настройки и эксплуатации.

4. Журналирование и отчеты (*A4*). В журналах должны аккумулироваться события и объекты, фиксируемые системой в ходе работы. Отчеты на основе собранной информации могут применяться для анализа действий пользователей, для расследования инцидентов информационной

безопасности, для контролясостояния защищенности информации, корректности настроек, и т.д.

5. Ценовая доступность (*A5*). Применяемые решения должны быть адекватными стоимости защищаемой информации и имеющимся финансовым возможностям компании.

6. Уровень ресурсоемкости (*A6*). Чаще всего заказчик настаивает на программном решении, которое легко внедряется в существующую инфраструктуру без изменения архитектуры сети, используемых прикладных программ и аппаратных средств.

Для комплексной оценки DLP-системы предлагается использовать алгоритм нечеткого вывода Мамдани, на первом шаге которого требуется сформировать базу правил системы нечеткого вывода [15].

Таким образом, входными параметрами базы правил системы нечеткого вывода являются шесть входных переменных *A1-A6*. Значения лингвистической переменной *A1* задаются с помощью терм-множества [«недостаточные», «достаточные»], а *A2-A6* — [«низкая», «средняя», «высокая»].

Выходной переменной базы правил системы нечеткого вывода является комплексная оценка DLP-системы, которая может представлять собой сложную нелинейную функцию от входных параметров.

База правил системы нечеткого вывода для оценки DLP-систем

База правил представляет собой конечное множество правил нечетких продукций, согласованных относительно используемых в них лингвистических переменных [16, 17].

С целью упрощения составления лингвистических правил лингвистические переменные были разбиты на две группы: эксплуатационную и функциональную. Эксплуатационная группа оценивается лингвистической переменной *B1*, функциональная группа — переменной *B2*. Значения

лингвистических переменных *B1* и *B2* задаются терм-множеством [«низкая», «средняя», «высокая»].

Для лингвистической переменной *B1* сформулированы 6 продукционных правил:

1. *Если A1=Недостаточные То B1=Низкая;*
2. *Если A2=Низкая То B1=Низкая;*
3. *Если A1=Достаточные И A2=Средняя И A6=Низкая То B1=Средняя;*
4. *Если A1=Достаточные И A2=Средняя И (A6=Средняя ИЛИ A6=Высокая) То B1=Низкая;*
5. *Если A1=Достаточные И A2=Высокая И (A6=Средняя ИЛИ A6=Низкая) То B1=Высокая;*
6. *Если A1=Достаточные И A2=Высокая И A6=Высокая То B1=Средняя,*

адля лингвистической переменной B2 — 10 продукционных правил:

1. *Если A4=Низкая То B2=Низкая;*
2. *Если A5=Низкая То B2=Низкая;*
3. *Если A3=Высокая И A4=Высокая То B2=Высокая;*
4. *Если A3=Высокая И A4=Высокая И (A5=Средняя ИЛИ A5=Высокая) То B2=Высокая;*
5. *Если A3=Средняя И A4=Средняя И A5=Средняя То B2=Средняя;*
6. *Если A3=Низкая И A4=Высокая И A5=Высокая То B2=Высокая;*
7. *Если A3=Низкая И A4=Средняя То B2=Низкая;*
8. *Если A3=Высокая И A4=Высокая И (A5=Средняя ИЛИ A5=Высокая) То B2=Высокая;*
9. *Если A3=Высокая И A4=Средняя И (A5=Средняя ИЛИ A5=Высокая) То B2=Средняя;*
10. *Если A3=Средняя И A4=Высокая И (A5=Средняя ИЛИ A5=Высокая)*

То B2=Высокая.

Значения выходной лингвистической переменной*R*, показывающей комплексную оценку, задаются терм-множеством [«низкая», «средняя», «высокая»]. Продукционные правила комплексной оценки DLP-системы выражены через лингвистические переменные *B1* и *B2*:

1. *Если B1=Низкая То R=Низкая;*
2. *Если B1=Средняя И B2=Низкая То R=Низкая;*
3. *Если B1=Средняя И (B2=Средняя ИЛИ B2=Высокая) То R=Средняя;*
4. *Если B1=Высокая И B2=Низкая То R=Средняя;*
5. *Если B1=Высокая И (B2=Средняя ИЛИ B2=Высокая) То R=Высокая.*

Однако в случае сложности составления правил и выбора функций принадлежности можно использовать адаптивную систему нейро-нечеткого вывода ANFIS (Adaptive Neuro-Fuzzy Inference System), реализованную в пакете Fuzzy Logic Toolbox системы MatLab [18]. С целью набора обучающей базы для сети ANFIS эксперты для различных комбинаций входных параметров определяют эффективность DLP-системы [19]. Затем происходит обучение системы ANFIS аналогично обучению искусственной нейронной сети. При этом осуществляется подбор, как функций принадлежности, так и продукционные правила.

Заключение

Использование предложенной базы правил в системе нечеткого вывода комплексной оценки DLP-систем, позволит аргументировано подобрать DLP-систему для удовлетворения нужд конкретной организации, сэкономить материальные ресурсы при закупке программного обеспечения, обеспечить построение эффективной системы информационной безопасности.

Литература

1. Papadimitriou P., Garcia-Molina H. Data Leakage Detection. URL: ilpubs.stanford.edu:8090/839/1/2008-23.pdf.

2. Айдинян А.Р., Цветкова О.Л., Кикоть И.Р., Казанцев А.В., Каплун В.В. О подходе к оценке информационной безопасности предприятия // Системный анализ, управление и обработка информации: сб. тр. V Междунар. науч. семинара, п. Дивноморское, 2-6 окт. Ростов н/Д: ДГТУ, 2014. С. 109-111.

3. Маршаков Д.В., Цветкова О.Л., Айдинян А.Р. Нейросетевая идентификация динамики манипулятора // Инженерный вестник Дона, 2011, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2011/504.

4. Кикоть И.Р., Айдинян А.Р., Цветкова О.Л. Методика выбора комплектующих для сельскохозяйственной техники на основе интеллектуальной системы // Состояние и перспективы развития сельскохозяйственного машиностроения: сб. ст. 8-й междунар. науч.-практ. конф., 3-6 марта.Ростов н/Д: ДГТУ, 2015. С. 296-298.

5. Айдинян А.Р., Цветкова О.Л. Методика оценки качества обучения студентов вуза с использованием нейро-нечеткого подхода // Программные продукты и системы, 2016. Т. 29. № 4. С. 189-193.

6. Чуйкова Е.Н. Реализация нечеткого выбора оборудования в системе проектирования информационной сети // Вестник ДГТУ, 2014. Т. 14, № 3 (78). С. 164-171.

7. Темичев А.А., Файзрахманов Р.А. Подбор параметров нагрузочного тестирования систем мониторинга с использованием нечеткой логики // Инженерный вестник Дона, 2015, № 3. URL: iv-don.ru/ru/magazine/archive/n3y2015/3153.

8. Айдинян А.Р., Цветкова О.Л. Генетические алгоритмы распределения работ // Вестник ДГТУ, 2011. Т. 11, № 5 (56).С. 723-729.

9. Айдинян А.Р., Цветкова О.Л. Алгоритм доставки сельскохозяйственной продукции несколькими исполнителями // Состояние и перспективы развития сельскохозяйственного машиностроения: сб. ст. 6-й междунар. науч.-практ. конф., 26 февр.-1 марта. Ростов н/Д, 2013. С. 247-248.

10. Айдинян А.Р., Цветкова О.Л., Панасенко Н.Д., Воронков И.В. Алгоритм формирования последовательности доставки грузов несколькими исполнителями // Математические методы в технике и технологиях: сб. тр. XXVII междунар. науч. конф., 3-5 июня. Тамбов. 2014, № 1 (60). С. 53-55.

11. Жолобова Е.А., Жолобова О.А., Гинеева А.В., Шаповаленко Я.И., Бакулин А.В. Метод комплексной оценки инвестиционных проектов в строительстве на основе нечеткой логики // Инженерный вестник Дона, 2017, № 1. URL: ivdon.ru/ru/magazine/archive/n1y2017/4059.

12. Гридневский А.В. Комплексная оценка геологических опасностей территорий Ростовской области // Инженерный вестник Дона, 2013, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1946.

13. Цветкова О.Л., Айдинян А.Р. Интеллектуальная система оценки информационной безопасности предприятия от внутренних угроз // Вестник компьютерных и информационных технологий, 2014. № 8 (122). С. 48–53.

14. Song D. Practical Techniques for Searches on Encrypted Data / D. Song, D. Wagner, A. Perrig. URL: cs.berkeley.edu/~dawnsong/papers/se.pdf.

15. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д. Рудинского. М.: Горячая линия-Телеком, 2013. 384 с.

16. Чуйкова Е.Н., Заслонов С.А. Нечеткий поиск средств защиты растений в базе данных // Состояние и перспективы развития сельскохозяйственного машиностроения: сб. ст. 10-й междунар. юбилейной науч.-практ.конф. в рамках 20-й междунар. агропромышленной выставки «Интерагромаш-2017». Ростов н/Д, 2017. С. 340-342.

17. Кикоть И.Р., Чуйкова Е.Н. Анализ алгоритмических методов формирования термов лингвистических переменных // Молодой исследователь Дона. 2016. № 1. С. 22-27.

18. Штовба С.Д. Проектирование нечетких систем средствами MatLab. М.: Горячая линия-Телеком, 2007. 288 с.

19. Маршаков Д.В., Айдинян А.Р., Цветкова О.Л. Генерация обучающей выборки для нейросетевой модели технологических объектов и систем // Математические методы в технике и технологиях — ММТТ. Саратов: Саратовский государственный технический университет имени Гагарина Ю.А. 2014. № 2. С. 8-10.

References

1. Papadimitriou P., Garcia-Molina H. Data Leakage Detection. URL: ilpubs.stanford.edu:8090/839/1/2008-23.pdf.

2. Ajdinyan A.R., TSvetkova O.L., Kikot' I.R., Kazantsev A.V., Kaplun V.V. Sistemnyj analiz, upravlenie i obrabotka informatsii: sb. tr. V Mezhdunar. nauch. Seminara, Divnomorskoe, 2-6 okt., Rostov n/D: DGTU, 2014. pp. 109-111.

3. Marshakov D.V., TSvetkova O.L., Ajdinyan A.R. Inženernyj vestnik Dona (Rus), 2011, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2011/504.

4. Kikot' I.R., Ajdinyan A.R., TSvetkova O.L. Sostoyanie i perspektivy razvitiya sel'skokhozyajstvennogo mashinostroeniya: sb. st. 8-j mezhdunar. nauch.-prakt. konf., 3-6 marta, Rostov n/D: DGTU, 2015. pp. 296-298.

5. Ajdinyan A.R., TSvetkova O.L. Programmnye produkty i sistemy. 2016, № 4. pp. 189-193.

6. CHujkova E.N. Vestnik DGTU. 2014, vol. 14, № 3 (78). pp. 164-171.

7. Temichev A.A., Fajzrakhmanov R.A. Inženernyj vestnik Dona (Rus). 2015, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2015/3153.

8. Ajdinyan A.R., TSvetkova O.L. Vestnik DGTU. 2011, vol. 11, № 5 (56). pp. 723-729.

9. Ajdinyan A.R., TSvetkova O.L. Sostoyanie i perspektivy razvitiya sel'skokhozyajstvennogo mashinostroeniya: sb. st. 6-j mezhdunar. nauch.-prakt. konf., 26 fevr.-1 marta, Rostov n/D, 2013. pp. 247-248.

10. Ajdinyan A.R., TSvetkova O.L., Panasenko N.D., Voronkov I.V. Matematicheskie metody v tekhnike i tekhnologiyakh: sb. tr. XXVII mezhdunar. nauch. konf., 3-5 iyunya, Tambov, 2014, № 1 (60). pp. 53-55.

11. ZHolobova E.A., ZHolobova O.A., Gineeva A.V., SHapovalenko YA.I., Bakulin A.V. Inženernyj vestnik Dona (Rus), 2017, № 1. URL: iv-don.ru/ru/magazine/archive/n1y2017/4059.

12. Gridnevskij A.V. Inženernyj vestnik Dona (Rus), 2013, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1946.

13. TSvetkova O.L., Ajdinyan A.R. Vestnik komp'yuternykh i informatsionnykh tekhnologij. 2014, № 8 (122). pp. 48–53.

14. Song D. Practical Techniques for Searches on Encrypted Data. D. Song, D. Wagner, A. Perrig. URL: cs.berkeley.edu/~dawnsong/papers/se.pdf.

15. Rutkovskaya D., Pilin'skij M., Rutkovskij L. Nejronnye seti, geneticheskie algoritmy i nechetkie sistemy [Neural networks, genetic algorithms and fuzzy systems]: Per. s pol'sk. I.D. Rudinskogo. Moskow: Goryachaya liniya-Telekom, 2013. 384 p.

16. CHujkova E.N., Zaslonov S.A. Sostoyanie i perspektivy razvitiya sel'skokhozyajstvennogo mashinostroeniya: sb. st. 10-j mezhdunar. yubilejnoj nauch.-prakt. konf. v ramkakh 20-j mezhdunar. agropromyshlennoj vystavki «Interagromash-2017», Rostov n/D, 2017. pp. 340-342.

17. Kikot' I.R., CHujkova E.N. Molodoj issledovatel' Dona (Rus), 2016, № 1. pp. 22-27.

18. SHtovba S.D. Proektirovanie nechetkikh sistem sredstvami MatLab [Designing fuzzy systems using MatLab]. M.: Goryachaya liniya-Telekom, 2007. 288 p.

19. Marshakov D.V., Ajdinyan A.R., TSvetkova O.L. Matematicheskie metody v tekhnike i tekhnologiyakh. MMTT, Saratov: Saratovskij gosudarstvennyj tekhnicheskij universi-tet imeni Gagarina YU.A. 2014, № 2. pp. 8-10.