Труды XXVII научной конференции по радиофизике

СЕКЦИЯ
«ИНФОРМАЦИОННЫЕ СИСТЕМЫ. СРЕДСТВА, ТЕХНОЛОГИИ, БЕЗОПАСНОСТЬ»

Председатель – Л.Ю. Ротков, секретарь – А.А. Рябов.
Нижегородский государственный университет им. Н.И. Лобачевского.

ПРИМЕНЕНИЕ КВАНТОВОГО РАСПРЕДЕЛЕНИЯ КЛЮЧЕЙ ДЛЯ МОДИФИКАЦИИ КРИПТОГРАФИЧЕСКИХ ПРОТОКОЛОВ

Д.А. Глозштейн, И.Г. Сидоркина

Волгатех

Квантовое распределение ключей (QKD) использует фундаментальную физику для обеспечения безопасности. Несмотря на очевидные преимущества, всё ещё остается несколько препятствий для широкого применения. Например, потери в канале ограничивают как скорость передачи секретного ключа, так и расстояние передачи при разумном балансе между скоростью передачи и потерей данных. Для преодоления таких барьеров в последнее время стали популярны следующие два подхода: разработка квантовых ретрансляторов [1] и внедрение концепции доверенных узлов [2]. К сожалению, квантовые трансляторы требуют квантовой памяти с высокоточной запутанностью, которые коммерчески недоступны. С другой стороны, в реальных условиях трудно проверить степень доверия к передаваемым данным между любыми двумя узлами в оптической сети. QKD действительно может быть использован для построения будущих защищенных сетей. К сожалению, скорость генерации секретных ключей (SKR) для современных систем QKD с дискретной переменной (DV) очень низка, что затрудняет работу этих сетей. Системы с непрерывной переменной (CV)-QKD могут быть использованы для улучшения SKR, однако, как показано в новых исследованиях [3], достижимое расстояние передачи для схем CV-QKD значительно короче по сравнению со схемами с использованием полей-близнецов (TF) DV-QKD [4]. Постквантовая криптография (PQC) является рекомендуемой альтернативой QKD [5]. Тем не менее, нет доказательств того, что алгоритмы PQC не поддаются взлому сложными квантовыми алгоритмами [6].

Более того, очень часто криптография на решетках основана на хеш-функциях устойчивости к столкновениям, таких как: $u = Ax$, где x – секретный ключ, u – открытый ключ, $A – m*n$ общедоступная матрица, описывающая решетку. Злоумышленнику достаточно использовать эффективный алгоритм квантовой инверсии матрицы, аналогичный алгоритму Харроу-Хасидима-Ллойда (HHL) [7], чтобы определить секретный ключ отправителя по $u = A^{-1}x$ и, таким образом, взломать протокол PQC. Одним из вариантов для решения основных проблем как QKD, так и PQC, является их совместное использование. Однако, даже несмотря на то, что дальность передачи при использовании протокола TF-QKD [8] может быть удвоена, скорость передачи секретных ключей на порядки ниже скоростей передачи данных, используемых в современных системах оптической связи.

Поэтому более эффективным решением может быть использование другой стратегии для преодоления вышеуказанных проблем для QKD и PQC. Учитывая низкие значения скорости генерации ключей в протоколах QKD, можно использовать традиционные схемы QKD только на стадии инициализации протоколов. Ключевая идея состоит в том, чтобы использовать генерируемую в QKD последовательность не в качестве защищенного ключа, а для:

1) защиты последовательности открытых ключей;
2) задания начальных значения для соответствующих генераторов случайных хэш-функций отправителя и получателя;

3) определения параметров для инициализации протоколов; и т.д.

Эти защищенные последовательности будут намного короче длины ключа для одноразового шифрования, поэтому скорость генерации ключа соответствующего протокола QKD не будет серьезной проблемой. Безусловно, данный подход не решает проблему с ограниченным расстоянием передачи данных в протоколах QKD.

Для решения этой проблемы возможно использование различных стратегий:

1) совместное использование протоколов QKD и PQC;
2) использование низкоорбитальных спутников;
3) использование квантовых ретрансляторов с коррекцией ошибок.

QKD-улучшенные протоколы вычислительной безопасности

Предлагаемая концепция применима к любому вычислительному криптографическому протоколу. В качестве примера приведем модификацию для одного из наиболее популярных алгоритмов распределения открытых ключей, что сделает его устойчивым к атакам на основе квантовых компьютеров.

Сначала формализуем квантово-расширенное распределение открытых ключей. Чтобы инициализировать протокол, отправитель и получатель запускают QKD, чтобы получить общую последовательность больших целых чисел *{g}* и больших простых чисел *{n}*, а также общие начальные значения. Далее отправитель и получатель используют общие начальные значения для случайного выбора базового g и простого числа n, которые используются однократно. Отправитель случайным образом выбирает большое целое число x, рассчитывает $X = g^x \bmod n$ и отправляет X получателю. Получатель случайным образом выбирает большое целое число y, вычисляет $Y = g^y \bmod n$ и отправляет Y отправителю. Отправитель вычисляет ключ K_A по формуле: $K_A = Y^x \bmod n = g^{xy} \bmod n$. Получатель вычисляет ключ K_B по формуле: $K_B = X^y \bmod n = g^{xy} \bmod n$. Очевидно, что оба ключа идентичны, $K_A = K_B$. Поскольку отправитель и получатель используют случайно выбранную пару *{g, n}* только для одного ключа, чтобы взломать протокол злоумышленнику пришлось бы использовать метод полного перебора.

Оригинальный алгоритм шифрования Rivest-Shamir-Adleman (RSA) показан на рис. 1.

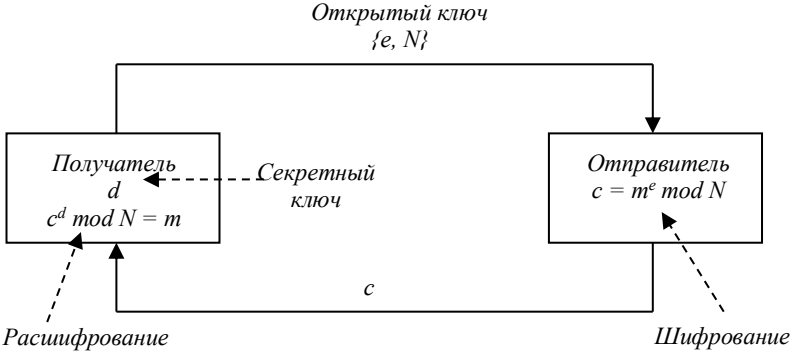


Рис. 1

Получатель случайным образом выбирает два простых числа p и q, чтобы получить $N = pq$. Он также выбирает e, у которого нет общего делителя с $(p - 1)(q - 1)$, в качестве

открытого ключа. Далее он вычисляет d как обратный элемент $e \bmod (p-1)(q-1)$ и использует его в качестве закрытого ключа. Далее он открыто отправляет $\{e, N\}$ отправителю. Чтобы отправить сообщение m отправитель шифрует с помощью $m^e \bmod N = c$ и отправляет c получателю. Он расшифровывает сообщение следующим образом: $c^d \bmod N = m$.

Чтобы взломать протокол, злоумышленнику нужно сначала определить период r функции $f(x)=m^x \bmod N = f(x+r)$ $(r=0,1,\ldots,2^n-1)$. Период функции $f(x)$ может быть найден на одном из шагов алгоритма факторизации Шора, который требует $O(n^3)$ элементарных операций $(2^n > N^2)$. Как только период r определен, злоумышленник может определить секретный ключ получателя, вычисляя $d' = e^{-1} \bmod r$ и взломать протокол RSA, определив переданное сообщение m следующим образом:

$$c^{d'} modN = (m^e)^{d'} modN = m^{ed'} modN = m^{1+kr} modN$$
$$= mm^{kr} modN = mmodN,$$

где $ed' = 1+kr$, $m^{kr} = 1 \bmod N$.

Необходимо модифицировать протокол RSA, как показано на рис. 2, чтобы квантовый компьютер не мог взломать его за полиномиальное время.

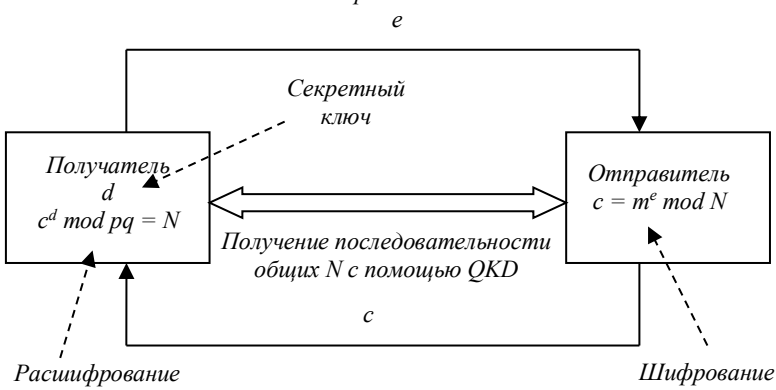


Рис. 2

Для этого можно инициализировать модифицированный алгоритм RSA, запустив протокол QKD, чтобы получить последовательность общих простых чисел $\{p\}$ и $\{q\}$, а также общие начальные значения. После инициализации отправитель и получатель будут использовать общие начальные значения для случайного выбора p и q, чтобы получить $N = pq$. Это N будет использовано однократно и немедленно уничтожено. Используя другое N для каждого нового ключа, злоумышленник не сможет определить его путем анализа зашифрованного текста, и будет вынужден применить метод прямого перебора.

[1] Duan L.-M., Lukin M. D., Cirac J. I., and Zoller P. Long-distance quantum communication with atomic ensembles and linear optics. Nature. 2001. Vol. 414, No. 6862. P. 413.

[2] Qiu J. Quantum communications leap out of the lab. Nature. 2014. Vol. 508. No. 7497. P. 441.
[3] Djordjevic B. Physical-Layer Security and Quantum Key Distribution. Cham, Switzerland: Springer Nature Switzerland AG. 2019.
[4] Lucamarini M., Yuan Z. L., Dynes J. F., Shields A. J. Overcoming the rate–distance limit of quantum key distribution without quantum repeaters. Nature. 2018. Vol. 557. P. 400.
[5] McEliece R. J. A public-key cryptosystem based on algebraic coding theory. DNS Progress Report. – Pasadena: Jet Propulsion Laboratory, 1978, p. 114.
[6] Micciancio D., Regev O. Lattice-based Cryptography. In D.J. Bernstein, J. Buchmann, E. Dahmen (eds) Post-Quantum Cryptography. – Berlin, Heidelberg: Springer, 2009.
[7] Harrow W., Hassidim A., Lloyd S. Quantum algorithm for solving linear systems of equations. Phys. Rev. Lett.2009. Vol. 103. P. 150502.
[8] Ma X., Zeng P., Zhou H. Phase-matching quantum key distribution. Phys. Rev. X. 2018. Vol. 8. P. 031043.

ПРИМЕНЕНИЕ ВЕСОВ ТЕРМОВ В ЗАДАЧЕ ОБНАРУЖЕНИЯ СПАМА

С.В. Корелов[1)], А.М. Петров[1)], И.Г. Сидоркина[2)], Л.Ю. Ротков[3)]

[1)] НКЦКИ
[2)] Волгатех
[3)] ННГУ им. Н.И. Лобачевского

Введение

Доказано, что спам является угрозой безопасности информации, нейтрализация которой является актуальной задачей. В связи с этим исследование, разработка, создание и внедрение новых и совершенствование существующих решений и технологий обеспечения безопасности информационных систем, ориентированных на обнаружение (выявление) спама, является актуальной и практически значимой задачей.

При этом создание новых моделей электронных писем, обеспечивающих выделение признаков электронных писем на основе их содержания с учетом меняющихся информационных потребностей конкретного пользователя, для обнаружения спама является актуальной задачей и представляет научный и практический интерес [1].

В связи с этим в [2] предложена и в [3] уточнена модель электронных писем, позволяющая специфическим способом (с помощью «генетических карт» [4]) выделять текстовые отрезки электронных писем, являющиеся отражением их отличительных признаков, или термы:

$$\Psi_{el} = \langle EL\_PrePoc, T\_Proc, T \rangle, \qquad (1)$$

где T_Proc – процедура выделения термов T – значимых последовательностей исходных символов текста электронного письма;

EL_PrePoc – процедура предобработки текста электронного письма.

Модель оперирует с преобразованными в числовой вектор данными, полученными из исходных текстов электронных писем. Корректность, практическая применимость и результативность модели (1) экспериментально продемонстрированы в [например, 2, 3].

Настоящая статья посвящена анализу результатов применения весовых коэффициентов термов в задаче обнаружения спама с использованием модели электронных писем (1).

Основная часть

Важной оставляющей в построении признаковых описаний текстов электронных писем является процедура расчета весов термов, которые существенно влияют на качество решения задачи их классификации [1]. В [1] продемонстрировано, что различные методы расчета весов термов дают в целом похожие результаты и различаются в зависимости от наборов данных. При этом наиболее широкое распространение имеют весовые функции класса $TF - IDF$ [например, 5, 6], как наиболее простые и показавшие свою относительно лучшую (в сравнении с другими) эффективность не только при решении задач классификации, но и поиска требуемой информации в массивах разнообразных текстов, а также их индексирования и рубрицирования.

Обозначим за tf_{ij} частоту -го терма в i-м письме – отношение числа его вхождений в текст письма к общему числу всех термов в этом письме. Тогда:

$$tf_{ij} = \frac{n_{ij}}{N_T}, \quad (2)$$

где n_{ij} – количество вхождений -го терма в i-м письме,

N_T – количество термов в -м письме.

Обозначим через idf_j инверсную документарную частоту -го терма – логарифм отношения числа всех писем к числу писем, в которых встречается j-й терм:

$$idf_j = log\left(\frac{N_{EL}}{n_j}\right), \quad (3)$$

где N_{EL} – количество всех писем,

n_j – количество писем, в которых встречается -й терм.

Основываясь на формулах (2) и (3) в базовом виде формулу расчета веса $TF - IDF$ можно представить следующим образом:

$$w_{ij} = tf_{ij} \cdot idf_j \text{ или} \quad (4)$$

$$w_{ij} = \frac{n_{ij}}{N_T} \cdot log\left(\frac{N_{EL}}{n_j}\right). \quad (5)$$

Особенностью данной меры является то, что вес терма пропорционален частоте его употребления в конкретном письме и обратно пропорционален частоте употребления во всех письмах. Таким образом, можно оценить важность терма в пределах конкретного письма. При этом больший вес получают термы с большей частотой в пределах конкретного письма и с меньшей частотой употребления в других письмах.

У формулы (5) существуют следующие наиболее распространенные модификации, которые целесообразно рассмотреть в качестве весов термов в задаче обнаружения спама с использованием модели электронных писем (1).

Очевидно, что $w_{ij} = 0$ при $N_{EL} = n_j$. Данный случай может наблюдаться при небольшом количестве классифицируемых писем. Во избежание таких случаев целесообразно применение [7] сглаживающего коэффициента:

$$w_{ij} = log\left(\frac{n_{ij}}{N_T} + 1\right) \cdot log\left(\frac{N_{EL}+1}{n_j}\right). \quad (6)$$

В [7] в дополнение к TF и IDF предложено ввести параметр CF (*аббр. от англ.* Class Frequency), обозначающую частоту терма в пределах заданного класса:

$$cf_j = \frac{N_{ij}^C}{N_i^C}, \quad (7)$$

где N_{ij}^C – количество писем того же класса, что и -е письмо, в которых встречается j-й терм.

N_i^C – число писем того же класса, что и -е письмо.

Тогда формула (6) примет вид:

$$w_{ij} = log\left(\frac{n_{ij}}{N_T} + 1\right) \cdot log\left(\frac{N_{EL}+1}{n_j}\right) \cdot \frac{N_{ij}^C}{N_i^C}. \quad (8)$$

В большинстве ситуаций в небольших письмах, как правило, будет присутствовать небольшое количество термов, а в больших наоборот. Это предопределяет необходимость использования коэффициента нормализации, позволяющего устранить эффект больших различий в частотах термов в текстах писем различной длины. В качестве коэффициента нормализации, как правило, может выступать следующий [7]:

$$norm_i = \frac{1}{\sqrt{\sum_{j=1}^{N_T}(w_{ij})^2}}. \quad (9)$$

С учетом коэффициента нормализации формулы расчета весов (5) и (6) примут соответственно вид:

$$w_{ij} = \frac{\frac{n_{ij}}{N_T} log\left(\frac{N_{EL}}{n_j}\right)}{\sqrt{\sum_{j=1}^{N_T}\left(\frac{n_{ij}}{N_T} log\left(\frac{N_{EL}}{n_j}\right)\right)^2}}, \quad (10)$$

$$w_{ij} = \frac{log\left(\frac{n_{ij}}{N_T}+1\right) \cdot log\left(\frac{N_{EL}+1}{n_j}\right)}{\sqrt{\sum_{j=1}^{N_T}\left(log\left(\frac{n_{ij}}{N_T}+1\right) \cdot log\left(\frac{N_{EL}+1}{n_j}\right)\right)^2}}. \quad (11)$$

В ряде исследований [например, 5, 8], посвященных вопросам анализа текстов и информационного поиска, приведен вариант меры $TF - IDF$ в формулировке поисковой системы INQUERY [9]:

$$w_{ij} = \beta + (1 - \beta) \cdot tf_{ij} \cdot idf_j, \quad (12)$$

где

$$tf_{ij} = \frac{tf_{ij}}{tf_{ij}+0{,}5+1{,}5 \cdot \frac{N_T}{\overline{N_T}}}, \quad (13)$$

$$idf_j = \frac{log\left(\frac{N_{EL}+0{,}5}{n_j}\right)}{log(N_{EL}+1)}, \quad (14)$$

где N_T– среднее число термов в одном письме (в термах);

$\beta = 0{,}4$ [5].

Для проведения эксперимента по применению весов термов в задаче обнаружения спама с использованием модели электронных писем (1) в качестве значений ее параметров приняты следующие:

$q = 256$ – соответствует количеству символов кодировки Windows-1251;

$\Delta t = 1$ – шаг дискретизации равен одному символу;

$n = 1$.

На основании полученных и обоснованных в [3] результатов все письма указанного англоязычного набора прошли следующую предварительную предобработку:

- удаление повторений символов пробелов, табуляции и переносов строк;
- перевод всех букв в верхний регистр.

Эксперимент и оценка его результатов проводились аналогично [3, 10, 11] на трех сформированных наборах электронных писем (Enron14, Enron 25 и Enron 36). При проведении экспериментальных исследований использован метод кросс-валидации [12].

В качестве мер оценки результатов эксперимента использованы точность *P*, полнота *R* обнаружения (классификации) и сбалансированная -мера [3, 10, 11, 13, 14]. Результаты эксперимента по каждой из этих мер представлены на рисунках 1-3.

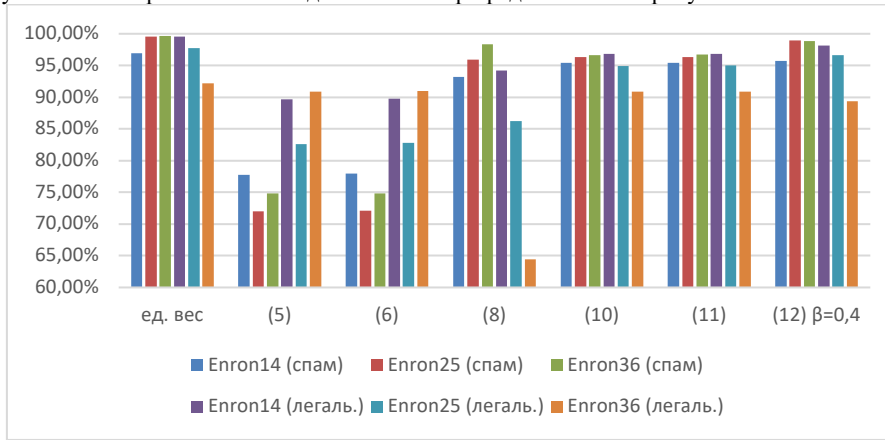


Рис. 1 – точность

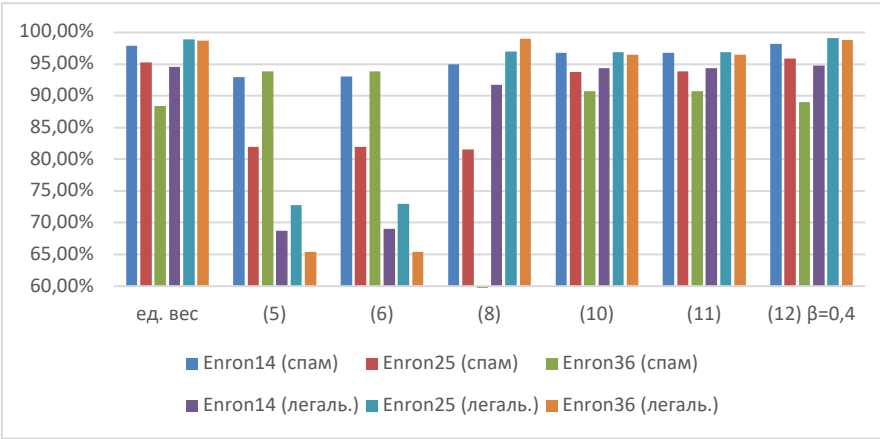


Рис. 2 – полнота

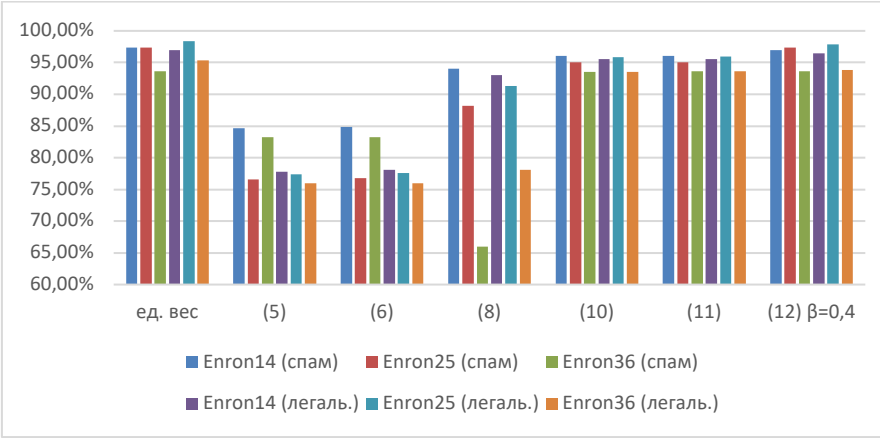


Рис. 3 – сбалансированная -мера

Анализ полученных результатов показывает, что наилучшие результаты обнаружения показывает единичный вес, а также вариант меры $TF - IDF$ (12) в формулировке поисковой системы INQUERY.

Заключение

Таким образом, результаты эксперимента свидетельствуют о целесообразности применения весовых коэффициентов термов, в том числе с целью исключения фактора случайности в процессе классификации, обусловленного возможным достижением максимального (предельного) числа уникальных термов в наборах соответствующих классов при бесконечном увеличении числа писем в обучающих наборах.

[1] Корелов С.В., Петров А.М., Сидоркина И.Г., Ротков Л.Ю. Применение весов термов в задаче обнаружения спама с использованием модели электронных писем // Труды XXVI научной конференции по радиофизике (Нижний Новгород, 12-27 мая 2022 г.). – Нижний Новгород: ННГУ, 2022. С. 522.

[2] Корелов С.В., Петров А.М., Ротков Л.Ю., Горбунов А.А. Модель электронных писем в задаче обнаружения спама // Вестник Поволжского государственного технологического университета. Сер.: Радиотехнические и инфокоммуникационные системы. 2020. № 2 (46). С. 44. DOI:10.25686/2306-2819.2020.2.44.

[3] Корелов С.В., Петров А.М., Ротков Л.Ю., Горбунов А.А. Предобработка текстов электронных писем в задаче обнаружения спама // Труды учебных заведений связи. 2020. Т. 6, № 4. С. 80. DOI:10.31854/1813-324X-2020-6-4-80-90.

[4] Кирьянов К.Г. Генетический код и тексты: динамические и информационные модели сложных систем /Ред. Л.Ю. Ротков, А.В. Якимов. – Нижний Новгород: ТАЛАМ, 2002, 100 с.

[5] Агеев М.С. Методы автоматической рубрикации текстов, основанные на машинном обучении и знаниях экспертов: дис. … канд. физ.-мат. наук: 05.13.11/Агеев Михаил Сергеевич. – Москва, 2004. – 136 с.
[6] Church K., Gale W. Inverse Document Frequency (IDF): A Measure of Deviations from Poisson // Natural Language Processing Using Very Large Corpora. 1999. Vol. 11. PP. 283-295. DOI:10.1007/978-94-017-2390-9_18.
[7] Liu M., Yang J. An Improvement of TFIDF Weighting in Text Categorization // 2012 International Conference on Computer Technology and Science (ICCTS 2012). 2012. Vol. 47. PP. 44-47. DOI:10.7763/IPCSIT.2012.V47.9.
[8] Лукашевич Н.В. Модели и методы автоматической обработки неструктурированной информации на основе базы знаний онтологического типа: дис. … докт. техн. наук: 05.25.05/Лукашевич Наталья Валентиновна. – Москва, 2014. – 312 с.
[9] Callan J.P., Croft W.B., Harding S.M. The INQUERY Retrieval System // Database and Expert Systems Applications. 1992. PP. 78-83. DOI:10.1007/978-3-7091-7557-6_14.
[10] Корелов С.В., Петров А.М., Ротков Л.Ю., Горбунов А.А. К вопросу об определении численного значения параметра в модели электронных писем // Труды XXIV научной конференции по радиофизике, посвященной 75-летию радиофизического факультета (Нижний Новгород, 13-31 мая 2020 г.). Нижний Новгород: ННГУ, 2020. С. 471-474.
[11] Корелов С.В., Петров А.М., Ротков Л.Ю., Горбунов А.А. Определение длины выборки в модели электронных писем // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. 2020. № 4 (36). С. 31-47. DOI:10.15593/2224-9397/2020.4.02.
[12] Перекрёстная проверка. [электронный ресурс]. // ВикипедиЯ. Режим доступа: https://ru.wikipedia.org/wiki/Перекрёстная_проверка, свободный (дата обращения: 07.10.2019).
[13] Корелов С.В., Петров А.М., Ротков Л.Ю., Горбунов А.А. Комбинирование значений параметра модели электронных писем // Материалы XII Международной Интернет-конференции молодых ученых, аспирантов и студентов «Инновационные технологии: теория, инструменты, практика» (16 ноября – 31 декабря 2020 г.). – Пермь: ПНИПУ. 2021. С. 448-455.
[14] Sebastiani F. Machine Learning in Automated Text Categorization // ACM Computing Surveys. 2002. Vol. 34, No. 1, 2002, PP. 1-47, DOI:10.1145/505282.505283.

ПРИМЕНЕНИЕ СЕМАНТИЧЕСКИХ МАСОК В ЗАДАЧЕ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

А.А. Коротышева, С.Н. Жуков

ННГУ им. Н.И. Лобачевского

Разработка отечественных решений для автоматической идентификации скрытых устройств при проверке оборудования и технических средств, способствует решению задач, регламентированных требованиями нормативных документов ФСБ России к порядку действий для спецпроверки.

Данная работа посвящена использованию семантических масок при идентификации информационных объектов на примере рентгеновских изображений. Рентгеновские изображения широко используются для визуализации внутренних структур объектов, если по внешним признакам невозможно точно идентифицировать объекты или их части.

Для сегментации объектов по рентгеновским изображениям возможно применение как алгоритмов компьютерного зрения, таких как алгоритм активных контуров [1] или алгоритм обнаружения границ [2], так и подходов, основанных на глубоком обучении, например, сверточных нейронных сетей [3].

Создание семантических масок в данной работе реализовано с помощью нейронных сетей с архитектурами U-Net [4] и Segment Anything Model (SAM) [5]. Обе модели способны к обобщению на новые объекты.

Для демонстрации семантических масок был использован набор данных с рентгеновскими изображениями элементов питания (ЭП) и проведено выделение значимых областей внутри ЭП, содержащих информацию о внутренних структурах, таких как сепараторы, стержни и прокладки. Каждая комбинация таких внутренних структур является уникальной и определяет отдельный тип (класс) ЭП. С помощью анализа информации, полученной путем такой сегментации, может быть проведена идентификация компонентов ЭП [6].

Примеры получаемых семантических масок объектов, где каждая маска выражает один из заданных классов внутренних структур ЭП, представлены на рисунке 1.

На рисунке 2 изображены результаты создания плотных карт сегментации двух объектов разных идентифицируемых классов. Плотные карты создаются путем объединения приведенных на рисунке 1 масок изображений по пикселям, где каждый пиксель относится к определенному классу.

Рис. 1

Рис. 2

Создание семантических масок при идентификации информационных объектов позволяет не только обнаруживать скрытые устройства и производить их классификацию, но и выявлять дефекты и несоответствия технических характеристик исследуемых объектов стандартам [7]. Использование предлагаемых алгоритмов сегментации и создания плотных карт объектов позволяет повысить точность идентификации на изображениях и анализа объектов в системах, получающих данные от рентгеновской установки. Набор алгоритмов может быть дополнен и адаптирован для применения в ранее разработанной интеллектуальной системе [6].

[1] Annangi P., Thiruvenkadam S., Raja A., et al. A region based active contour method for x-ray lung segmentation using prior shape and low level features // IEEE International Symposium on Biomedical Imaging: From Nano to Macro. – Rotterdam, Netherlands: IEEE, 2010. P. 892.
[2] Saad M. N., Muda Z., Ashaari N. S., Hamid H. A. Image segmentation for lung region in chest X-ray images using edge detection and morphology // IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014). – Penang, Malaysia: IEEE, 2014. P. 46.
[3] Chouhan V., Singh S.K., Khamparia A., et al. A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images // Applied Sciences. 2020. Vol. 10, No. 2. P. 559.

[4] O. Ronneberger, P. Fischer, T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS. 2015. Vol. 9351. P. 234-241.
[5] S. Roy, T.A. Kirillov, E. Mintun, [и др.]. Segment Anything // Computer Vision and Pattern Recognition. 2023. 30 pages.
[6] Блатов Р.И., Вострякова Е.А., Москвин А.С., Чупров Д.А., Егоров Ю.С., Коротышева А.А., Милов В.Р., Дубов М.С., Кербенева А.Ю. Программа для ЭВМ «Прототип интеллектуальной системы идентификации немаркированных элементов питания с использованием методов машинного обучения» // Свидетельство об официальной регистрации программы для ЭВМ № 2022663863 от 20.07.2022 г.
[7] ГОСТ Р МЭК 86-1-96. Батареи первичные. Часть 1. Общие положения. М.: ИПК Издательство стандартов. 1997. 43 с.

ИНТЕГРАЦИЯ СКАНЕРОВ УЯЗВИМОСТЕЙ В СИСТЕМЫ МОНИТОРИНГА РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

М.М. Мунтян, И.Г. Сидоркина

ФГБОУ ВО «ЧГУ им. И.Н. Ульянова»

Введение

В настоящее время одним из наиболее популярных способов осуществления контроля состояния инфраструктуры организаций является программный мониторинг всех его компонент. Одними из средств, подходящих для этих целей являются системы мониторинга рисков информационной безопасности и сканеры уязвимостей. Однако оба решения обладают определенного рода недостатками, которые затрудняют проведение контроля с точки зрения информационной безопасности. Например, системы мониторинга в своем большинстве ориентированы на накопление информации об инфраструктуре и осуществление контроля за действиями пользователей, в том числе в части определения эффективности работы сотрудников. В то же время сканеры уязвимостей более приспособлены для решения вопросов, связанных с обеспечением безопасности, однако проведения определенного рода проверок способны нанести ущерб функционированию сканируемых объектов.

С целью снижения негативного влияния, оказываемого обоими продуктами на функционирование инфраструктуры, а также организации их нормального использования, предлагается произвести интеграцию сканеров уязвимостей с системами мониторинга рисков информационной безопасности.

Сканеры уязвимостей

Под сканерами уязвимостей понимаются программные или аппаратные инструментальные средства, основная задача которых заключается в сборе информации о защищенности информационной инфраструктуры в режиме реального времени, путем проведения «простого» сканирования или более сложных процедур, имитирующих реальные пути использования уязвимостей [1]. Основной причиной их востребованности является наличие возможности определения неправомерных действий со стороны потенциальных нарушителей или неквалифицированные действия собственных сотрудников организации.

Системы мониторинга

Системы мониторинга рисков информационной безопасности – это программное и аппаратное обеспечение, которое создано для того, чтобы производить выявление и фиксацию событий, связанных с нарушением информационной безопасности в инфраструктуре организации. Однако, если основная роль сканеров заключается в выявлении уязвимостей инфраструктуры, главная цель систем мониторинга заключается в определении и анализе любых событий, связанный с ней (информационной безопасности).

Рассуждая в таком ключе логично утверждать, что сами по себе системы мониторинга являются тем же сканерами уязвимостей, однако ориентированными не только на сами уязвимости, но и на другие события информационной безопасности [2, 3].

Интеграция сканеров уязвимостей в системы мониторинга рисков информационной безопасности.

Предложено решение по интеграции сканеров уязвимостей и систем мониторинга рисков информационной безопасности должно осуществляться в соответствии со следующими принципами (см. рис.).

- Введение сканера уязвимостей информационной безопасности как одно из штатных механизмов системы мониторинга (элемент программного агента системы мониторинга), который при этом позволял бы проводить автономное сканирование без передачи информация на сервер системы мониторинга для ситуаций эталонного сканирования (под эталонным сканированием подразумевается проведения сканирования с целью проверки устранения уже определенных ранее уязвимостей).
- Предоставление сканерам уязвимостей возможности первичной обработки информации при проведении мониторинга в результате которого удастся произвести первичную фильтрацию информации и построения подходящего шаблона уязвимости информационной безопасности на основе ее основных и дополнительных признаков.
- Организация взаимодействия по соотнесению выявленных уязвимостей и угроз информационной безопасности на основе данных хранящихся на сервере системы мониторинга.

Конечный продукт состоит из следующих элементов.

Программный агент – децентрализованный элемент системы мониторинга, обладающий частичной автономией, который производит обмен информации с серверами системы мониторинга в асинхронном режиме (установление соединения по необходимости без учета деятельности других программных агентов).

Сервер обработки информации (сервера баз данных) и центр принятия решений (аналитический элемент, сервера для передачи повторных запросов) – элемент системы мониторинга способный производит накопление информации, полученной от объектов, а также способный инициировать запросы к другим агентам с целью сокращения времени анализа безопасности инфраструктуры в целом. Одной из задач этого элемента является проведения соотнесения информации, полученной ранее от программных агентов, о связях между уязвимостями и угрозами информационной безопасности. Данный элемент должен быть реализован как совокупность таких элементов, как сервер системы мониторинга и ПЭВМ (группы ПЭВМ) – аналитических агентов, производящих последующую обработку, полученных от программных агентов информации.

При реализации аналитических элементов целесообразным является использование методов интеллектуального анализа, как механизмов совершенной обработки больших объемов информации, которые являются невозможными для человека.

При реализации серверных элементов необходимо формирование базы данных таким образом, чтобы представленная в ней информация сохранялась в полном виде без сокращений. Такой подход позволит производить проверку решений и эффективности работы продукта, а также выявлять неточности и ошибки путем добавление в процессы определения угроз человека.

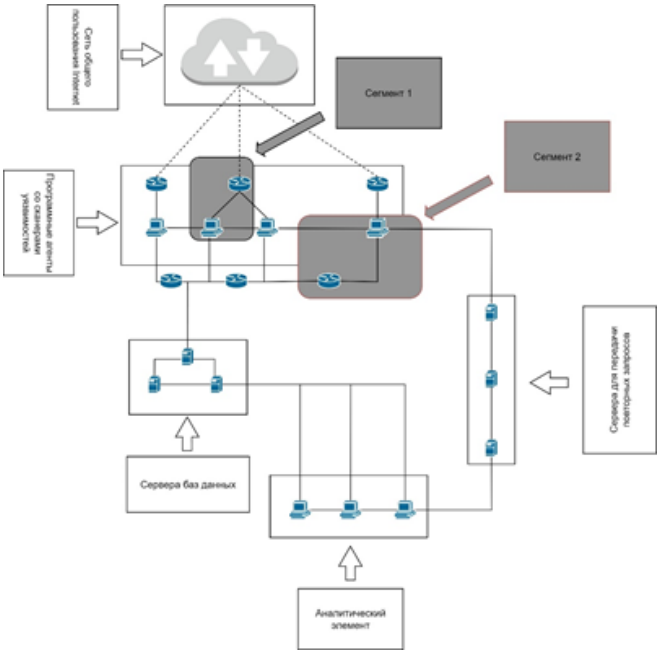


Рис.

Необходимость интеграции сканеров уязвимостей и систем мониторинга предложенным способом связана с такими важными факторами обеспечения безопасности информации как время, вероятность совершения ошибки при проведении анализа и объемом информации для него.

Под временем понимается такой период, в течении которого инфраструктуры является практически беззащитной для реализации угроз информационной безопасности. Прежде всего это связано с тем, что используемым механизмам защиты необходимо адаптироваться к происходящему, а также произвести идентификацию инцидента, а также ресурсов, которые были подвергнуты атаке. В результате чего централизованное управление всеми элементами инфраструктуры является достаточно затратным относительно этого параметра, что делает его менее эффективным относительно децентрализованного функционирования, когда каждый из элементов может осуществлять необходимые действия независимо от других. Однако использование децентрализованного управления требует жесткого разграничения областей деятельности каждого элемента (сегментация), что также необходимо учитывать при настройке предлагаемого решения.

Факторы ошибок несовершенных средств автоматизации и объемов анализируемых данных являются взаимосвязанными. Это обуславливается тем, что даже при отдельном использовании системы мониторинга организациям необходимо учитывать

объемы работы по анализу данных, которые в свою очередь достаточно велики. В результате чего задача становится емкой относительно ресурсов, которые необходимы использовать для ее решения, при этом такое решение не гарантирует отсутствие ошибок, так как процесс соотнесения является субъективным и означает присутствие вероятностных факторов в ее решении. На этом основано существование компаний, предоставляющих услуги по круглосуточному мониторингу инцидентов информационной безопасности (компаний, оказывающих услуги SOC – Security Operations Center). Иными словами, количество записей в системном журнале и иных средствах получения информации для систем мониторинга настолько велико, что даже выделенное для этого подразделение организации или компания оказывающая услуги на основе договора будет вынуждена потратить достаточно длительный промежуток времени для проведения идентификации инцидента и выработки ответных на него действий. При этом подобная процедура не гарантирует достоверность полученных результатов.

Таким образом, для успешной интеграции сканеров уязвимостей с системами мониторинга рисков информационной безопасности необходимо произвести модернизацию продуктов в соответствии с представленными принципами. В то же время необходимо включить в рассмотрение добавление в функционал систем мониторинга методов интеллектуального анализа, что сведет участие человека в процессе к минимуму, а также позволит производить вычисления намного точнее и быстрее.

Заключение

Таким образом, предложенное решение позволяет произвести ориентацию систем мониторинга рисков информационной безопасности на основные пути реализации угроз безопасности информации – уязвимые элементы инфраструктуры, через которые эти угрозы могут быть реализованы. В результате чего процесс обеспечения безопасности становится схожим с процессом оказания медицинской помощи, где для обеспечения безопасности пациентов производится лечение не симптомов заболевания, а причин их возникновения, что благоприятно влияет на состояние инфраструктуры в целом. Такой вывод вытекает исходя из того, что решение проблем уязвимостей путем их устранения или осуществления контроля над ними, позволяет производить решение возникающих инцидентов намного эффективнее.

В это же время сканеры уязвимостей при такой интеграции освобождаются от необходимости осуществления полного сканирования инфраструктуры из одной точки, что в значительной степени снижало эффективность их работы. При этом способе интеграции сканеры функционируют независимо друг от друга в рамках своего сегмента, что значительно снижает время проведения сканирования и способствует снижению рисков для функционирования других сегментов инфраструктуры.

[1] https://www.reg.ru/blog/chto-takoe-skanery-uyazvimostej/
[2] https://lib.itsec.ru/articles2/control/monitoring_analiz_upravl_ib;
[3] https://rtmtech.ru/articles/monitoring-informatsionnoj-bezopasnosti/;

ПРИМЕНЕНИЕ ЗАКОНА БЕНФОРДА В ОБНАРУЖЕНИИ СГЕНЕРИРОВАННЫХ ИЗОБРАЖЕНИЙ

С.П. Никитенкова

ННГУ им. Н.И. Лобачевского

Прогресс нейросетей GAN (Generative adversarial network) позволил значительно улучшить качество синтетических изображений или дипфейков. С одной стороны, это может использоваться в таких областях, как реклама, кинопроизводство, видеоигры и т.д. С другой стороны, дипфейки представляют собой серьезную угрозу для общества, политической системы и бизнеса. Миллионы сгенерированных цифровых изображений загружаются каждый день в интернет, распространяясь в социальных сетях, размывая границы между фактами и вымыслом. Изображения искусственно сгенерированных лиц, способны обмануть даже самых опытных наблюдателей, и, главное, вызвать симпатию и большее доверие, чем настоящие лица [1].

Закон Бенфорда широко используется в мультимедийной криминалистике для обнаружения фальсификации изображений. Как оказалось, интенсивность света на реальных изображениях при определенных ограничениях точно подчиняется закону Бенфорда, что позволяет обнаружить применение фильтров и ретушь изображений.

Соответствие закону Бенфорда квантованных коэффициентов дискретного косинусного преобразования изображений, сжатых в формате JPEG, неоднократно демонстрировалось как мощный инструмент для обнаружения манипуляций с изображениями, в том числе в качестве индикатора морфинга изображения лица.

Закон Бенфорда, также известный как закон первой цифры или закон значащей цифры, является эмпирическим законом. Справедливость закона Бенфорда была продемонстрирована и подтверждена в различных областях. Примерами являются распределение результатов выборов, суммарная длительность нот в классических музыкальных произведениях, фальсификация научных данных, данные заболеваемости и смертности от коронавирусной инфекции и т.д.

Впервые закон был обнаружен Ньюкомбом в 1881 г. и переоткрыт Бенфордом в 1938 году. В законе говорится, что распределение вероятностей первых цифр x (x = 1, 2, ..., 9) в наборе натуральных чисел является логарифмическим:

$$P(d) = log_{10}\left(1 + \frac{1}{d}\right).$$

Спектральные свойства изображения можно проанализировать с помощью дискретного преобразования Фурье. Для дискретного двумерного сигнала $f(x, y)$, представляющего отдельные цветовые каналы изображения размера $M \times N$, дискретное преобразование Фурье $F(k_x,k_y)$ определяется как:

$$F\left(k_x, k_y\right) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) exp\left(-2\pi i\left(\frac{k_x x}{M} + \frac{k_y y}{N}\right)\right),$$

где x, y – позиция пикселя, $f(x,y)$ – значение пикселя выбранного канала изображения, k_x,k_y – пространственная частота.

Спектр мощности изображения является важной статистической характеристикой изображения и определяется как

$$P(k_x, k_y) = \left|F(k_x, k_y)\right|^2.$$

После применения преобразования Фурье к изображению информация представлена в новой области, но по-прежнему содержит 2D-информацию. Размерность может быть уменьшена без существенной потери информации путем азимутального усреднения:

$$P(r) = \frac{1}{2\pi}\int_0^{2\pi} P(r,\theta)d\theta, \text{ где } r = \sqrt{\frac{k_x^2+k_y^2}{\frac{1}{4}(M^2+N^2)}} \text{ и } \theta = atan2(k_x, k_y).$$

Генерация изображений чаще всего происходит с помощью методов, основанных на так называемых генеративно-состязательных сетях, сокращенно GAN (Generative adversarial network). Впервые представленные в 2014 году, сети GAN завоевали популярность благодаря своей способности создавать фотореалистичные изображения с нуля. Технология StyleGAN (Style Generative Adversarial Network) является расширением архитектуры GAN. Сети StyleGAN обеспечивает генерацию изображений на основе стилей, что позволяет контролировать синтез генерируемых изображений. Первый вариант технологии StyleGAN был опубликован в 2019 году. В 2020 году была предложена технология StyleGAN2, позволившая добиться значительного улучшения качества изображений. В октябре 2021 года компанией NVIDIA была опубликована архитектура StyleGAN3 (Alias-free), главной целью которой стала адаптация технологии StyleGAN для применения в анимации и видео.

Чтобы проверить, было ли распределение значений азимутально-усредненного спектра мощности изображения, и значений, полученных в соответствие с законом Бенфорда, равным или нет, был рассчитаны критерии согласия, основанные на тестах Крамера-Мизеса и Коломогорова-Смирнова.

Тесты измеряют расстояние между наблюдаемыми и ожидаемыми в соответствие с законом Бенфорда значениями. Критерий Крамера-Мизеса рассчитывался как:

$$W^2 = \frac{1}{N}\sum_{i=1}^{9}(S_i - T_i)^2 t_i,$$

где $S_i = \sum_{j=1}^{i} q_j$, и $T_i = \sum_{j=1}^{i} p_j$, обозначают совокупные наблюдаемые и ожидаемые величины. Величина t_i определяется как $t_i=(p_i+p_{i+1})/2$ (i=0, 1..8) и $t_9=(p_9+p_1)/2$.

Критерий Колмогорова-Смирнова вычисляет как:

$$KS = \sqrt{N}\ \sup_{1\le i\le 9}|S_i - T_i|.$$

Были рассмотрены датасеты FFHQ, StyleGAN2, StyleGAN3, каждый из N=100 изображений [2-4]. Для каждого изображения была выдвинуты гипотезы: подчиняется ли случайная величина значений азимутально-усредненного спектра мощности изображения распределению Бенфорда (нулевая гипотеза) или нет. Количество случаев, когда нулевая гипотеза верна, для каждого датасета представлено в таблице.

Табл.

	Число изображений (нулевая гипотеза верна)					
	Cramér–von Mises test			Kolmogorov-Smirnov test		
	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
Датасет FFHQ	70	48	22	83	83	30
Датасет StyleGAN2	51	27	6	76	32	6
Датасет StyleGAN3	68	42	17	85	85	19

Результаты показывают, что изображения, сгенерированные StyleGAN2, часто не соответствуют закону Бенфорда, на этом основании их можно отличить от реальных изображений. Однако изображения, сгенерированные StyleGAN3, имеют практически тот же процент соответствия, что и реальные изображения на исследуемых датасетах. Изображение, сгенерированные StyleGAN3, труднее обнаружить предложенным способом, что мотивирует дальнейшие исследования по этой теме.

[1] Europol (2022), Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union, Luxembourg https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf
[2] GitHub – NVlabs/FFHQ-dataset. https://github.com/NVlabs/ffhq-dataset
[3] GitHub – NVlabs/StyleGAN2-dataset. https://github.com/NVlabs/stylegan2
[4] https://www.kaggle.com/datasets/showmik50/stylegan3-dataset

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ В РАМКАХ ПОСТРОЕНИЯ SIEM И ДРУГИХ СИСТЕМ ЗАЩИТЫ ИНФОРМАЦИИ

Л.М. Плотников, Р.Г. Нужный, Л.Ю. Ротков, В.А. Мокляков

ННГУ им. Н.И. Лобачевского

Введение

Неизбежная зависимость от цифровых процессов на объектах КИИ означает, что отказ систем может привести к огромным потерям ресурсов и эффективности, а возможно и к человеческим жертвам. Принимая во внимание уже существующие угрозы со стороны «недружественных» государств, вопрос защиты объектов КИИ и линий связи стоит сейчас особенно остро.

На сегодняшний день атаки на информационную инфраструктуру детектируются с помощью специального оборудования и средств, при этом специалистами по-прежнему используются «ручные» инструменты для сканирования сетевого трафика в том числе на предмет подозрительной активности, или при расследовании компьютерных инцидентов. Так или иначе, в случае неоднозначного детектирования атаки, ложного срабатывания, или нового вида атаки, решение о блокировке того или иного потока данных в инфраструктуре и выводе из эксплуатации зараженных машин принимает ответственный специалист по информационной безопасности на объекте. Применение методов машинного обучения (MO/ML) способны сделать работу этих специалистов ещё более эффективной. Если бы модель MO могла эффективно идентифицировать аномальные пакеты данных, проходящих через сеть ОКИИ, то специалисты по безопасности смогли бы тратить меньше времени на ручной анализ. Но для реализации такой модели MO, необходимо снизить уровень ложно-положительных и ложно-отрицательных срабатываний. Модель ML также должна быть эффективной и гибкой для того, чтобы работать с разными видами трафика и оборудования. В рамках данного тезиса была предпринята попытка построить прототип такой модели, обученной на уже доступном наборе данных.

Теория обнаружения аномалий

Обнаружение аномалий связано с поиском точек, которые отклоняются от большинства данных относительно их среднего или медианы в распределении. В машинном обучении ещё часто используется словосочетание "обнаружение выбросов" [1]. Некоторые модели обнаружения аномалий работают с одним признаком (одномерные данные), например, при мониторинге электрических сигналов. Однако большинство моделей обнаружения аномалий используют многомерные данные, что означает наличие двух (двумерных) или более (многомерных) признаков.

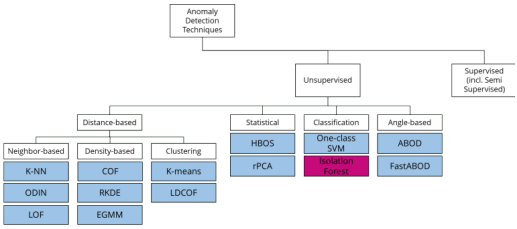


Рис. 1

Обнаружение выбросов – это проблема классификации. Однако эта область более разнообразна, поскольку обнаружение выбросов – это проблема, к которой можно подойти с помощью методов обучения с учителем и без [2]. На рисунке 1 представлена диаграмма, которая даёт хороший обзор стандартных алгоритмов, которые обучаются без надзора. Необходимым условием для контролируемого обучения является наличие информации о том, какие точки данных являются выбросами, а какие относятся к обычным данным. От того, известно ли, какие классы в наборе данных являются выбросами, а какие нет, зависит выбор возможных алгоритмов, которые мы могли бы использовать для решения проблемы обнаружения выбросов. Методы обучения без надзора являются естественным выбором, если метки классов недоступны. Если же метки классов доступны, можно использовать как алгоритмы обучения без надзора, так и алгоритмы обучения с надзором.

Поиск набора данных

При подготовке к реализации модели рассматривалось несколько вариантов получения нужных наборов данных. Непосредственно перед экспериментом стоял вопрос использования промышленного TCP/IP трафика реальных SCADA-систем, однако доступа к нему получить не удалось. Была предпринята попытка смоделировать нужные данные, но специализированные генераторы, которые были найдены, более не предоставляются в России на коммерческой основе, не говоря уже об использовании в научных целях. Более простые, не позволяют симулировать атаку на систему (разве что кроме Dos-атак). Тогда методом поиска был найден набор данных отлично подходящий для поставленной задачи [3]. Необработанные сетевые пакеты набора данных UNSW-NB 15 были созданы в лаборатории Cyber Range Lab UNSW Canberra для генерации гибрида обычных действий и искусственного поведения современных атак. Было захвачено 100 ГБ необработанного трафика (например, файлы Pcap) [4, 5]. Часть из этих данных была обработана тренировочный набор с 175000 записями. Этот набор данных включает девять типов атак, а так же 49 признаков с меткой, например: время жизни пакета, количество переданных байт, категорию атак.

Описание алгоритмов

В рамках построения модели были выбраны алгоритмы: изолирующий лес, алгоритм k-ближайших соседей, алгоритм локального уровня выброса. Кратко рассмотрим их по порядку. Сутью алгоритма изолирующего леса является то, что аномальные точки данных легче отделить от остальной выборки. Для того, чтобы изолировать точку данных, алгоритм рекурсивно создаёт разделы выборки путём случайного выбора признака, а затем случайного выбора значения разделения между минимальным и максимальным значениями, допустимыми для этого признака.

Рис. 2

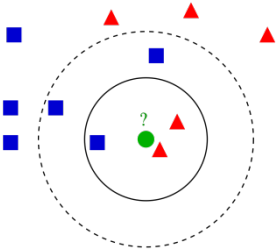


Рис. 3

Алгоритм k-ближайших соседей используется для классификации и регрессии. В обоих случаях входные данные состоят из k ближайших обучающих примеров в наборе данных. Результат зависит от того, используется ли алгоритм для классификации или регрессии. При классификации, результатом является принадлежность к классу. Объект классифицируется путём множественного голосования его соседей, при этом объект относится к классу, наиболее распространённому среди k ближайших соседей (k – целое положительное число, обычно небольшое). Если k = 1, то объекту просто присваивается класс единственного ближайшего соседа.

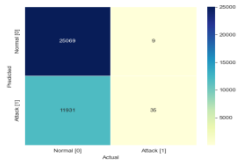
Алгоритм локального уровня выброса основывается на концепции локальной плотности, где локальность задаётся k ближайшими соседями, расстояния до которых используются для оценки плотности. Путём сравнения локальной плотности объекта с локальной плотностью его соседей можно выделить области с аналогичной плотностью и точки, которые имеют существенно меньшую плотность, чем её соседи. Эти точки считаются выбросами. Локальная плотность оценивается типичным расстоянием, с которым точка может быть «достигнута» от соседних точек. Определение «расстояния достижимости», используемого в алгоритме, является дополнительной мерой для получения более устойчивых результатов внутри кластеров.

Построение модели

Программа создавалась на языке программирования “Python”, с использованием библиотек: numpy, pandas, scikit-learn, matplotlib. На первом этапе был проанализирован тренировочный набор данных. Всего в нём 175000 записей, из которых почти половина записей искусственных атак на систему. Было принято решение сосредоточиться на одной категории атак. После этого, набор данных был разделён, перемешан и отнормирован. Для оценки алгиритмов использовались стандартные метрики в МО[6,7]: Precision, Recall, Accuracy, F1-score.

Изолирующий лес

Для тренировки были выбраны базовая модель, нужная для понимания, и улучшенная, в которой можно было менять количество деревьев, ожидаемую пропорцию выбросов к данным и т.д. (рис. 4).



Isolation Forest (baseline) model

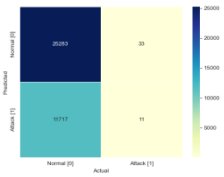
	precision	recall	f1-score	support
0	0.68	1.00	0.81	25078
1	0.80	0.00	0.01	11966
accuracy			0.68	37044
macro avg	0.74	0.50	0.41	37044
weighted avg	0.72	0.68	0.55	37044

f1_score: 40.67%

Рис. 4

Алгоритм локального уровня выброса

Для тренировки была выбрана улучшенная модель, изменения схожи с улучшенной моделью изолирующего леса (рис. 5).



LOF model

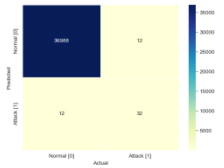
	precision	recall	f1-score	support
0	0.68	1.00	0.81	25316
1	0.25	0.00	0.00	11728
accuracy			0.68	37044
macro avg	0.47	0.50	0.41	37044
weighted avg	0.55	0.68	0.56	37044

f1_score: 40.67%

Рис. 5

Алгоритм k-ближайших соседей

Для тренировки были выбраны базовая модель, нужная для понимания, где количество соседей определялось автоматически и улучшенная, в которой можно было менять количество соседей для больших возможностей (рис. 6).



KNN (baseline) model

	precision	recall	f1-score	support
0	1.00	1.00	1.00	37000
1	0.73	0.73	0.73	44
accuracy			1.00	37044
macro avg	0.86	0.86	0.86	37044
weighted avg	1.00	1.00	1.00	37044

f1_score: 86.35%

Рис. 6

Сравнение результатов работы всех алгоритмов представлено на рисунке 7.

Рис. 7

Выводы

Учитывая, что изолирующий лес является самым популярным алгоритмом в таком типе задач, в данном случае итоги его работы нельзя назвать удовлетворительными: алгоритм плохо работает с большим их количеством и большим количеством признаков. Так же для работы алгоритма опасно большое количество аномалий, ибо их скопление он может посчитать за норму.

Алгоритм локального уровня выброса тоже неудовлетворительно себя проявил. У алгоритма мягкая метрика, поэтому бывает трудно интерпретировать конечные результаты. Так же локальность метода наверняка сыграла свою роль.

Алгоритм k-ближайших соседей же показал себя отлично, почти все метрики около 0.8. Конечно, по количеству правильно определённых угроз он уступает изолирующему лесу, но у того было огромное количество ложно-положительных результатов.

В целом очевидно, что причиной неудовлетворительной работы стали данные, а точнее неэффективный выбор и обработка признаков. Для лучшей работы алгоритмов придётся либо брать другой готовый набор данных, либо выбрать признаки и обработать данные из нынешнего набора ещё более тщательно. Но даже в таком виде, хотя бы 1 алгоритм из 3 идентифицировал большую часть угроз с минимальным количеством ложно-положительных срабатываний.

[1] Hodge V. J., Austin J. A Survey of Outlier Detection Methodologies // Artificial Intelligence Review. 2004. Vol. 22, No. 2. doi:10.1007/s10462-004-4304-y.
[2] Нужный Р.Г., Ротков Л.Ю., Мокляков В.А. Практическое применение классификатора сетевого трафика на основе методов машинного обучения Труды двадцать пятой научной конференции по радиофизике. – Н. Новгород: Изд-во ННГУ, 2021.
[3] Bierbrauer D., Chang A., Kritzer W., Bastian N. 2021. Cybersecurity Anomaly Detection in Adversarial Environments. https://doi.org/10.48550/arXiv.2105.06742
[4] Moustafa N., Slay J. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)."Military Communications and Information Systems Conference (MilCIS), 2015. IEEE. 2015.
[5] Moustafa N., Slay J. "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set." // Information Security Journal: A Global Perspective. 2016. P. 1.
[6] https://www.relataly.com/measuring-classification-performance-in-machine-learning-with-python-and-scikit-learn/846/
[7] David M. W. Powers. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlationю" // Journal of Machine Learning Technologies. 2011. Vol. 2, No. 1. P. 37.

МОДИФИЦИРОВАННАЯ ПРОЦЕДУРА УСТАНОВЛЕНИЯ СОЕДИНЕНИЯ В ВИРТУАЛЬНОЙ ЧАСТНОЙ СЕТИ

В.Д. Зюзин, А.А. Рябов

ННГУ им. Н.И. Лобачевского

В настоящее время технология виртуальных частных сетей (ВЧС) широко используется для организации безопасного обмена информацией как на уровне Intranet, так и на уровне Extranet. Чаще всего ВЧС строятся на базе семейства протоколов IPSec.

Рассмотрим ВЧС, содержащую N шлюзов.

Установление соединения по технологии IPSec происходит с использованием протокола Internet Key Exchange (IKE). Протокол IKE состоит из двух этапов. На первом этапе создается вспомогательный защищенный туннель между инициатором и конечным шлюзом. По этому туннелю будет проходить согласование параметров защищенного туннеля. В согласовании параметров защищенного туннеля состоит суть второго этапа [1]. В этой схеме уязвимое место – создание вспомогательного туннеля на первом этапе протокола IKE.

В работе рассматривается модифицированная процедура создания туннеля передачи данных.

В отличии от стандартной процедуры создания туннеля передачи данных, в модифицированной процедуре на первом этапе создаются три промежуточных вспомогательных туннеля. В том числе, туннель от шлюза-инициатора до произвольного промежуточного шлюза из числа N, входящего в состав ВЧС, от этого промежуточного шлюза до другого произвольного промежуточного шлюза, также входящего в состав ВЧС, и от последнего до оконечного шлюза.

На втором этапе происходит согласование параметров безопасности туннеля между инициатором и оконечным узлом. Согласование происходит по защищенному соединению, составленному из промежуточных туннелей, созданных на первом этапе.

После согласования параметров начинает работать туннель между конечными узлами, а соединения со вспомогательными узлами прекращаются по тайм-ауту.

Такая схема позволяет уменьшить вероятность проведения атаки «Человек по середине» в момент начала установления соединения, а значит увеличить защищенность установки соединения.

[1] Кульгин М. Технологии корпоративных сетей. Энциклопедия. — СПб.: Питер, 2000, 704 с

ПРОКТОРИНГ КАК СИСТЕМА ВЫЯВЛЕНИЯ АНОМАЛЬНОГО ПОВЕДЕНИЯ СТУДЕНТОВ ПРИ ПРОХОЖДЕНИИ ИНТЕРНЕТ-ТЕСТИРОВАНИЙ

Д.А. Семенов, И.Г. Сидоркина

Волгатех

На протяжении последних двух лет все большую актуальность обретает возможность дистанционного прохождения тестирования. При этом основное требование к сеансу тестирования остается прежним – самостоятельное выполнение экзаменационной работы тестируемым без посторонней помощи. Для обеспечения контроля за выполнением данного требования при дистанционной работе применяются системы прокторинга. Прокторинг – это процедура контроля при интернет-тестировании, при которой за всем процессом наблюдает администратор – проктор. Он следит за действиями экзаменуемого с помощью веб-камеры и видит, что происходит на экране компьютера тестируемого. Эта технология позволяет подтвердить личность кандидата, исключить списывание, помощь посторонних лиц, использование запрещенных ресурсов и прочие нарушения на экзамене. Кроме того, внедрение прокторинга позволяет проходить тестирование людям с ограниченными возможностями здоровья.

Актуальность исследования заключается в том, что с распространением данной технологии, применение человека-проктора становится все более трудоемким. С ростом числа одновременных сеансов тестирования требуется кратное увеличение числа прокторов. В связи с чем на данный момент разрабатываются системы автоматизированного прокторинга, где вместо человека за экзаменуемым наблюдает система на основе искусственного интеллекта. Это позволяет существенно снизить количество человеческих ресурсов при проведении сеансов тестирования, а также обеспечить непредвзятую и объективную оценку выполнения работы экзаменуемым.

Цель работы – расширение возможностей системы прокторинга платформе i-exam.ru посредством анализа позадачной статистики выполнения заданий тестируемыми.

Основная функциональность системы прокторинга обеспечивает трансляцию с экрана и веб-камеры, что позволяет отслеживать самостоятельное прохождение тестирования студентом. Дополнительно обеспечивается трансляция с камеры телефона, находящегося сбоку от тестируемого с целью максимального охвата рабочей зоны.

На данный момент система прокторинга на платформе i-exam.ru включает в себя следующие модули:

1) модуль проверки технических требований для проверки соответствия рабочего места тестируемого еще до начала экзамена во избежание проблем непосредственно во время тестирования. Проверка включает в себя запуск системы в поддерживаемом браузере, наличие веб-камеры и качество трансляции, возможность трансляции рабочего стола, отслеживание необходимой пропускной способности интернет-канала, подтверждение работоспособности протокола WebRTC;
2) модуль трансляции для передачи видео- и аудиоданных на WebRTC сервер на платформе i-exam.ru, а также остановкой тестирования в случае прерывания трансляции;
3) модуль записи, обеспечивающий сохранение трансляции;

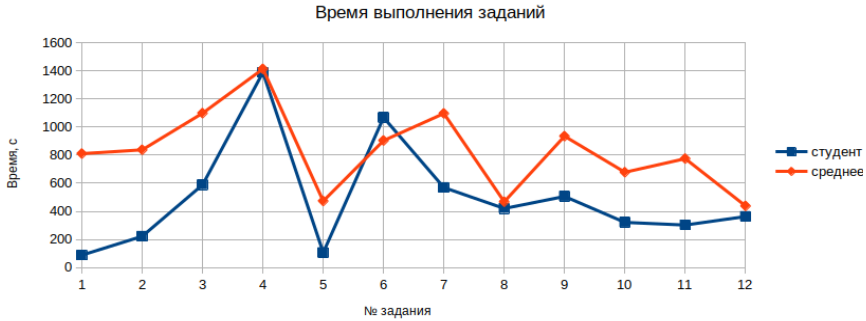
4) модуль выявления нарушений, обрабатывающий сохраненные на предыдущем этапе данные и передающий информацию о найденных нарушениях проктору;
5) модуль просмотра нарушений для проверки нарушений и исключения ложных срабатываний, по результатам просмотра принимается решение об аннулировании результата тестирования.

В автоматическом режиме (с применением искусственного интеллекта) возможно выявление таких явных нарушений регламента, как покидание рабочего места и появление посторонних лиц в кадре путем анализа изображения с камеры.

Тем не менее, одно только изображение с камеры и трансляция экрана монитора не всегда гарантируют самостоятельность выполнения работы тестируемым, в связи с чем разработка данной системы подразумевает анализ множества дополнительных факторов: затраченного времени на выполнение задания и тестирование в целом, набранного балла и т. д. В связи с этим, при проведении интернет-тестирования необходим сбор статистики, которая включает данные о затраченном времени и результате выполнения каждого задания сеанса тестирования. При сопоставлении этих данных всех пользователей имеется возможность выявить аномалии, такие как подозрительно быстрое или, напротив, медленное и правильное выполнение задания, существенное превышение набранного балла по сравнению с результатами других пользователей, неравномерность ответов на задания схожей тематики и др.

Предлагаемое решение заключается в разработке алгоритмов и методов выявления отклонения статистических характеристик решения задания от среднего. В частности, аномально быстрое решение статистически сложного задания может говорить о потенциальном факте несамостоятельного выполнения задания тестируемым. В таком случае проктору предоставляется подробный протокол выполнения задания с фиксацией видеозаписи, на основании которого будет сделан соответствующий вывод о самостоятельности выполнения задания тестируемым.

В качестве примера на рисунке приведена позадачная статистика выполнения заданий тестируемым и среднего времени выполнения заданий остальными участниками. В ходе анализа графика можно сделать вывод, что время выполнения некоторых заданий существенно ниже времени выполнения заданий другими участниками, что может говорить о недобросовестном прохождении сеанса тестирования.



Помимо анализа времени решения заданий, предлагается анализ коэффициентов решаемости каждого задания. Коэффициент решаемости задания представляет собой отношение количества верно данных ответов на задание к общему количеству ответов на задание. Отклонение коэффициента решаемости от среднестатистического может сигнализировать о необходимости дополнительной проверки результатов тестирования студента проктором.

Также предполагается анализ идентичности векторов ответов студентов. Схожие вектора ответов могут говорить о заимствовании ответов между тестируемыми.

В качестве примера в таблице представлены вектора идентификаторов ответов на задания двух студентов. Верно данные ответы отмечены зеленым цветом, неверные – красным.

Табл.

№ задания	1	2	3	4	5	6	7	8	9	10	11	12
Ответы студента 1	3	2	1	5	3	1	2	4	3	2	2	5
Ответы студента 2	3	2	1	7	3	1	2	3	3	2	2	5

Визуальная оценка позволяет сделать вывод, что студенты на большинстве заданий совершали схожие ошибки, что говорит о высокой вероятности заимствования ответов. Автоматизированным способом это может быть выявлено путем расчета коэффициента корреляции между двумя векторами.

Коэффициент корреляции (R) - количественная оценка тесноты взаимосвязи двух случайных величин. Значение коэффициента корреляции находится в диапазоне от 0 до 1. Абсолютное значение коэффициента корреляции показывает силу взаимосвязи элементов векторов. Чем выше значение модуля коэффициента корреляции, тем сильнее связь между элементами векторов.

Для данных векторов коэффициент корреляции составляет:

$$R = \sqrt{1 - \frac{\sum(y_i - y_x)^2}{\sum(y_i - \bar{y})^2}} = \sqrt{1 - \frac{3{,}57}{31{,}67}} = 0{,}942.$$

Полученная величина свидетельствует о том, что фактор x (ответы студента 1) существенно влияет на y (ответы студента 2), что говорит о недобросовестности выполнения работы одним из студентов.

Указанные выше способы анализа данных о тестировании в совокупности с информацией о нарушениях от системы прокторинга могут дать максимально полную и практически безошибочную информацию о достоверности сеанса тестирования.

[1] Семенов Д.А. // Инженерные кадры – будущее инновационной экономики России: материалы VIII Всероссийской студенческой конференции; Йошкар- Ола, 8-11 ноября 2022 года. – Йошкар-Ола: Поволжский государственный технологический университет, 2022. С. 465.

[2] Прокторинг в онлайн-экзаменах: как это работает? [Электронный ресурс] – Режим доступа: https://habr.com/ru/companies/stepic/articles/329420/

[3] Гмурман В. Е. Теория вероятностей и математическая статистика: Учебное пособие для вузов. — 10-е издание, стереотипное. — Москва: Высшая школа, 2004. — 479 с.

АЛГОРИТМ ГЕНЕРАЦИИ РЕЧЕПОДОБНОЙ ПОМЕХИ С ИДЕНТИФИКАЦИЕЙ ГОЛОСА ДИКТОРА ДЛЯ ЗАЩИТЫ КОНФИДЕНЦИАЛЬНЫХ ПЕРЕГОВОРОВ

Р.А. Васильев

ННГУ им. Н.И. Лобачевского

Общие положения

Защита акустической информации, циркулирующей в защищаемом помещении, входит в один из базисов мероприятий по информационной безопасности предприятия (организации, фирмы). Данные мероприятия реализуются с применением пассивных и активных методов защиты [1].

Пассивные методы защиты строятся на основе снижения вероятности получения информации или ее расшифровки из акустических источников с использованием различного вида звукопоглощающих материалов.

Активные методы защиты акустической информации - это методы, которые включают использование специального оборудования или программного аппаратных комплексов, для предотвращения утечки акустической информации. Активные методы основаны на создании дополнительных помех, которые скрывают сигнал, несущий речевую информацию, в каналах, где может быть утечка. В качестве маскирующих сигналов широко используется «белый» или «розовый» шум в диапазоне частот от 100 до 10000 Гц [2].

В последнее время начали применять комбинированные сигналы, включающие так называемые речеподобные сигналы [3-6].

Устройства активной защиты речевой информации, как правило, состоят из генератора маскирующих сигналов и набора преобразователей электрических сигналов в акустические или преобразователей электрических сигналов в механические перемещения.

Проведенные исследования показали, что наиболее эффективным является речеподобная помеха, формируемая из речевых сигналов. В статье предложен алгоритм формирования речеподобной помехи, представляющей собой случайную последовательность звуков речи с возможностью идентификации голоса диктора с применением метода обеляющего фильтра [7]. Формирование речеподобной помехи реализовано с использованием программного средства разработки MatLab в разработанной автором «Программе идентификации дикторов по голосу» (ИС ИДГ) [8], модернизированной для решения задачи генерации речеподобной помехи диктора. Эффективность предлагаемой речеподобной помехи оценена экспериментально.

Теоретический анализ

На данный момент времени специалистами предлагается три типа формирования речеподобной помехи (РЧП) [9].

В [10] основным показателем эффективности защиты речевого сигнала выбрана словесная разборчивость речи W_c.

Словесная разборчивость W_c показывает насколько понятна для оператора технических систем перехвата информации, очищенный речевой сигнал от систем защиты акустической информации.

Спектр речи разбивают на N октавных полос. Чаще всего используют среднегеометрические частоты в диапазоне от 125 Гц до 80000 Гц.

Исследования показывают, что при «Wc» менее:

- «50% – 70%» – невозможно полностью восстановить информационную составляющую разговора;
- «20 % – 40 %» – невозможно установить тему разговора;
- «20 %» – факт ведения разговор становится под вопросом.

Главная идея предложенного в статье алгоритма формирования РЧП с возможностью идентификации голоса диктора заключается не только в снижении коэффициента словесной разборчивости W_c, используемого для расчёта выполнения норм по противодействию речевой разведке при проведении конфиденциальных переговоров, но и значительное затруднение проведения цифровой шумоочистки перехваченного речевого сигнала, так как для генерации помехового сигнала используется не «белый шум», а РЧП с фонемами говорящего на совещании диктора.

Экспериментальные исследования

Для экспериментальных исследований была ИС ИДГ [11-13], модернизированная к задачам генерации РЧП диктора, посредством доработки модуля идентификации диктора по голосу. На рисунке изображен реализованный в ИС ИДГ алгоритм генерации речеподобных помех с идентификацией диктора по голосу.

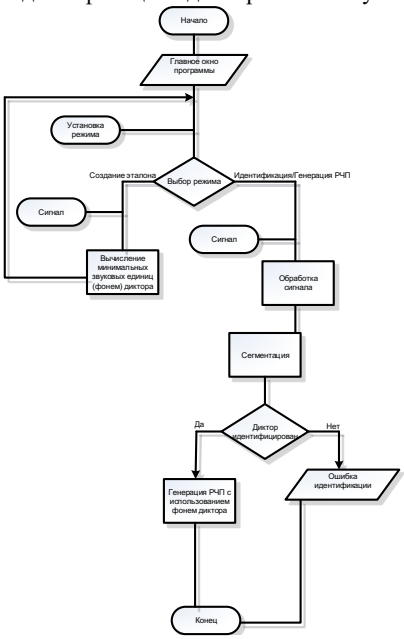


Рис.

В ходе эксперимента создана база фонем 10-ти дикторов. Произведено измерение акустического сигнала и расчёт коэффициента речевой разборчивости W_c по методике, описанной в [14], для трех случаев:

1) без применения средства акустической защиты (САЗ);
2) с применением САЗ, генерирующего помеху «белый шум»;
3) с применением генерации РЧП в дополнении к САЗ, генерирующему помеху «белый шум».

Эксперимент показал, что добавление РЧП снизило разборчивость с 28,3% до 7%.

Кроме этого была произведена шумоочистка для сигналов, записанных в случаях 2 и 3. Для случая 2 удалось разобрать защищаемую речь. Для случая 3 – не удалось.

[1] Хорев А. А. Способы защиты выделенных помещений от утечки речевой (акустической) информации по техническим каналам: системы виброакустической защиты // Специальная техника. – М.: 2013, № 4, с. 31.
[2] Хорев А. А. Системы виброакустической маскировки // Специальная техника. 2003. № 6. С. 28.
[3] Дворянкин С.В., Уленгов С.В., Устинов Р.А., Дворянкин Н.С., Антипенко А.О. Системное моделирование речеподобных сигналов и его применение в сфере безопасности, связи и управления // Безопасность информационных технологий. 2019. Т. 26, № 4. С. 101.
[4] Авдеев В.Б., Трушин В.А., Кунгуров М.А. Унифицированная речеподобная помеха для средств активной защиты речевой информации // Тр. СПИИРАН. 2020. Выпуск 19. Т. 5. С. 991.
[5] Хорев А.А., Царев Н.В. Способ и алгоритм формирования речеподобной помехи // Вестник ВГУ, серия: Системный анализ и информационные технологии. 2017. № 1. С. 57.
[6] Воробьев В.И., Давыдов А.Г. Синтез речеподобных сигналов // Акустический журнал, 2002. № 5. Т. 48. С. 701.
[7] Савченко В. В. Информационная теория восприятия речи. // Изв. вузов. Радиоэлектроника. 2007. Вып. 6. С. 3.
[8] Васильев Р. А. Свид. о гос. регистрации программы для ЭВМ №2015663306 Программа идентификации дикторов по голосу / Васильев Р.А. Зарег. 15.12.2015г. – М.: Роспатент, 2015.
[9] Хорев А.А. Безопасность информационных технологий [Электронный ресурс]. 2008. Режим доступа: http://www.security.ukrnet.net/d-book-2/ch_10.pdf. Дата доступа: 28.04.2023.
[10] Покровский Н. Б. Расчет и измерение разборчивости речи. – М.: Гос. изд-во лит. по вопросам связи и радио, 1962, 391 с.
[11] Васильев Р.А., Ротков Л.Ю. Адаптация метода биометрической идентификации по голосу к тихому произнесению парольных фраз для противодействия утечки речевой информации по акустическим каналам // Труды 25 Научной конференции по радиофизике – Н. Новгород: Изд-во ННГУ, 2021. С. 517.

[12] Васильев Р.А. Исследование фонетического строя речи и идентификация дикторов по голосу // Безопасность информационных технологий. 2013. Т. 20. № 1. С. 85-86.
[13] Васильев Р.А. Исследование особенностей фонетического строя речи и текстонезависимая идентификация дикторов по непрерывной речи // Информационная безопасность регионов. 2012. № 2 (11). С. 57-63.
[14] Васильев Р.А., Ротков Л.Ю. Оценка защищенности речевой информации от утечки по акустическим и виброакустическим каналам с помощью программно-аппаратного комплекса «Шёпот» // Учебно-методическое пособие, ННГУ, 2020. 57 С.

АКТИВНО-ПАССИВНЫЙ МЕТОД ОПРЕДЕЛЕНИЯ КООРДИНАТ ЦЕЛЕЙ В РЛС, СОПРЯЖЕННОЙ С ПАССИВНЫМ КАНАЛОМ РАДИОКОНТРОЛЯ ИЗЛУЧЕНИЙ СИСТЕМ РАДИОЛОКАЦИОННОГО ОПОЗНАВАНИЯ

И.Н. Карельский, Л.Ю. Ротков, Р.А. Верхов

ННГУ им. Н.И. Лобачевского

Одним из путей повышения эффективности современных активных РЛС, как основного источника получения информации о воздушных целях (ВЦ), является их интеграция со средствами пассивной локации и создание активно-пассивных комплексов (АПК), функционирующих как единая информационная система [1].

В качестве *объекта исследования* далее рассматривается АПК, включающий РЛС и размещаемое на одной позиции с ней пассивное средство радиотехнического контроля (СРТК) излучений систем радиолокационного опознавания (СРЛО) (рис 1).

Целью исследования является разработка *активно-пассивного угломерно-суммарно-дальномерного метода* (АПМ) определения плоскостных координат ВЦ, оборудованных бортовыми ответчиками (БО) СРЛО, в интересах повышения информативности, информационной безопасности и экономии энергетического ресурса РЛС.

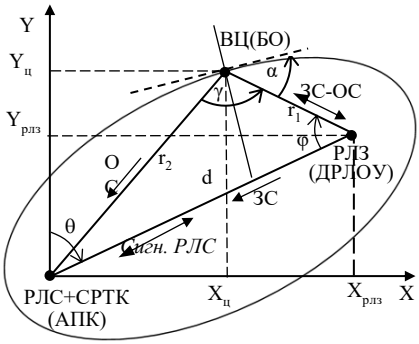


Рис. 1

Наблюдаемая в АПК СРЛО состоит из радиолокационного запросчика (РЛЗ), как самостоятельного стационарного объекта или узла (блока), размещаемого на какомлибо подвижном носителе (самолете, корабле и др.), и БО, устанавливаемого на дружественной для РЛЗ воздушной цели – ВЦ(БО). Примером такой системы является СРЛО ВВС США и НАТО MkXII, которая опознает ВЦ в режиме «запрос-ответ» с помощью периодических кодированных запросных и ответных импульсных радиосигналов (ЗС и ОС) [2].

Обычно РЛЗ сопрягается с обзорной РЛС и синхронно с ней совершает круговой или секторный периодический обзор пространства, излучая ЗС на несущей частоте $f_з$ через остронаправленную в горизонтальной плоскости антенну. При совпадении диаграммы направленности антенны (ДНА) РЛЗ с направлением на ВЦ(БО) ЗС поступает в БО и спустя малое детерминированное время срабатывания ($t_{фо}$) БО излучает на частоте $f_о$ ОС, по которому в РЛЗ происходит процедура опознавания.

Параметры излучаемых СРЛО MkXII ЗС и ОС достаточно известны [2] для того, чтобы организовать их прием двухканальным по частоте ($f_з, f_о$) радиоприемником СРТК. Расчеты показывают, что при чувствительности приемника не хуже -110 дБ/Вт дальность приема ЗС и ОС относительно РЛЗ и ВЦ(БО) будет ограничиваться только дальностью прямой видимости (ДПВ). Наибольшая дальность приема ЗС будет при размещении РЛЗ на воздушном носителе, например, на самолете дальнего радиолока-

ционного обнаружения и управления (ДРЛОУ) Boeing E-3 AWACS. При патрулировании AWACS на высоте 6...9 км указанная дальность может достигать 400 км. С другой стороны, дальность радиолокационного обнаружения самолета типа Boeing-707 (E-3) с эффективной поверхностью рассеяния 40…60 м^2 зависит от типа РЛС в АПК и также может достигать ДПВ [3], что свидетельствует о наличии общей для СРТК и РЛС зоны обнаружения носителей аппаратуры СРЛО.

В [4] рассмотрен пример реализации чисто пассивного угломерно-суммарно-дальномерного метода применительно к удаленному стационарному пункту контроля систем вторичной радиолокации с априорно известными координатами опорного РЛЗ. Однако в условиях конфликта доступность координат РЛЗ противника не очевидна.

Сущность метода. В АПМ в качестве опорного РЛЗ предлагается использовать стационарный или движущийся радиолокационно видимый РЛЗ, координаты которого контролируются с помощью РЛС. С этой целью СРТК первоначально пеленгует РЛЗ по принимаемым ЗС, например, используя моноимпульсную антенну, подобную антенне РЛС «Крона» [5]. Затем по пеленгу, полученному от СРТК, РЛС обнаруживает и определяет координаты РЛЗ в полярной (азимут θ, дальность d) и декартовой ($X_{\text{рлз}} = dsin\theta, Y_{\text{рлз}} = dcos\theta$) системах. Далее прием ЗС от обнаруженного РЛЗ происходит через изотропную антенну СРТК, в те моменты, когда ДНА РЛЗ находится на линии РЛЗ-СРТК.

Координаты ВЦ(БО) ($X_{\text{ц}}$, $Y_{\text{ц}}$) определяются в СРТК как точка пересечения двух линий положения (ЛП) постоянного значения измеряемых параметров.

Первая ЛП – *линия постоянного угла* φ (прямая «РЛЗ-БО»), определяемого по формуле: $\varphi = 2\pi \text{T}_{\text{зо}}/\text{T}_{\text{вращ}}$, где: $\text{T}_{\text{зо}}$ – интервал времени, измеренный между моментом прихода ЗС (ДНА РЛЗ на линии «РЛЗ-СРТК»), и моментом прихода ОС (ДНА РЛЗ на линии «РЛЗ-БО»); $\text{T}_{\text{вращ}}$ – предварительно измеренный в СРТК период вращения антенны РЛЗ по двум эпизодам прихода в СРТК ЗС.

Вторая ЛП – *линия постоянной суммы расстояний* $R_\Sigma = r_1 + r_2$ (r_1 – «РЛЗ-БО» и r_2 – «БО-СРТК»), являющаяся эллипсом с фокусами в точках РЛЗ и СРТК. Величина R_Σ определяется по времени запаздывания очередного поступившего в СРТК импульса ОС ($t_{\text{о}}$) относительно времени очередного импульса ЗС ($t_{\text{з}}$) с учетом времени $t_{\text{фо}}$: $R_\Sigma = \text{c}(t_{\text{о}} - t_{\text{з}} - t_{\text{фо}})$. Поскольку в момент $t_{\text{о}}$ ДНА РЛЗ направлена на БО и текущий запросный импульс на входе приемника СРТК отсутствует, то момент его отсчета прогнозируется по предварительно измеренному в СРТК периоду запроса.

Измеренные величины φ, R_Σ, d позволяют рассчитать расстояние «РЛЗ-БО» [3]:

$$r_1 = \frac{{R_\Sigma}^2 + 2R_\Sigma d}{2R_\Sigma + 2dcos\varphi}. \qquad (1)$$

Тогда выражения для вычисления координат цели будут иметь вид:

$$X_{\text{ц}} = X_{\text{рлз}}\, r_1 \sin(180 - \theta \pm \varphi)\,; Y_{\text{ц}} = Y_{\text{рлз}}\, r_1\, cos(180 - \theta \pm \varphi) \qquad (2)$$

В выражениях (2) знак (-) перед φ ставится, если вращение антенны ДНА РЛЗ происходит против хода часовой стрелки; знак (+), если вращение по ходу часовой стрелки.

Оценка точности метода. Для оценки точности определения координат ВЦ(БО) предварительно определяются среднеквадратические ошибки (СКВО) определения

двух ЛП [6] в которых должны быть учтены ошибки определения координат РЛЗ в РЛС (σ_θ, σ_d).

СКВО прямой: $\sigma_y = \sigma_\varphi r_1$, где σ_φ – СКВО измерения угла φ, зависящая от ширины ДНА РЛЗ ($2\varphi_{0,5}$). Для моноимпульсной ДНА: $\sigma_\varphi = (0{,}02 \div 0{,}03)2\varphi_{0,5}$ [7]. С учетом ошибки измерения в РЛС базы d (σ_d) можно записать $\sigma_y = ((\sigma_\varphi r_1)^2 + (\sigma_d)^2)^{1/2}$.

СКВО эллипса: ($\sigma_э$=$\sigma_{R_\Sigma}/2\cos(\gamma/2)$, где: σ_{R_Σ} – СКВО измерения R_Σ, γ – угол, под которым видна суммарно-дальномерная база d из точки ВЦ(БО). Угол γ при известных параметрах треугольника АПК-ВЦ(БО)-РЛЗ определяется по формуле $\gamma = \arcsin(dsin\varphi/R_\Sigma - r_1)$. Величину σ_{R_Σ} с учетом ошибки измерения пеленга РЛЗ в РЛС (σ_θ) можно определить по формуле $\sigma_{R_\Sigma} = (\sigma_э^2 + \sigma_о^2 + (\sigma_\theta d)^2)^{1/2}$, где $\sigma_э$ и $\sigma_о$ СКВО ошибок измерения расстояния по временному положению импульсов ЗС и ОС. Они определяются их длительностями ($\tau_э, \tau_о$) и величиной отношения сигнал/шум q [7]: $\sigma_о = c\tau_0/\sqrt{(\pi q)}$; $\sigma_э = c\tau_э/\sqrt{(\pi q)}$.

Тогда СКВО координат цели $\sigma_ц$ при независимых измерениях ЛП можно найти, как ошибку определения точки пересечения двух ЛП [6]:

$$\sigma_ц = (\sigma_э^2 + \sigma_y^2)^{1/2}/sin\alpha, \qquad (3)$$

где α=(π-γ)/2 – угол под которым пересекаются линии положения.

Знание ошибок $\sigma_ц$ в различных точках окружающего АПК пространства позволяет выбрать рабочую зону реализации АПМ, где эти ошибки не превышают максимального значения ($\sigma_ц \le \sigma_{ц\,max}$), задаваемого потребителем информации АПК.

На рис. 2 отображено поле ошибок (СКВО) АПМ в виде окружностей радиуса r=$\sigma_ц$ на участке пространства (400×400 км). Для наглядности размер окружностей укрупнен в масштабе (4:1).

В качестве исходных данных для расчета ошибок использованы: $\tau_э$=0,8 мкс; τ_0=0,45 мкс; q=20 дБ; $t_{фо}$=3 мкс; $2\varphi_{0,5} = 1^\circ$; σ_θ=25′; σ_d=100 м.

На рис. 3 показана зависимость величины ошибки от расстояния АПК-РЛЗ (величины базы) при различном удалении РЛЗ-ВЦ(БО).

Выводы:

1. Ошибки определения координат ВЦ(БО) предложенным АПМ в пределах совместной рабочей зоны обзорной РЛС и СРТК соизмеримы или незначительно превышают ошибки активного метода локации, поэтому полученные в АПК координаты целей могут использоваться в качестве основного или дополнительного целеуказания различным потребителями информации, например, огневым средствами ПВО.

2. Метод позволяет снизить до минимума время работы РЛС на излучение и сопровождать ВЦ(БО) в режиме повышенной временной скрытности, что затруднит ведение радиотехнической разведки РЛС бортовыми средствами этих целей и улучшит её живучесть в условиях применения самонаводящегося на излучение оружия.

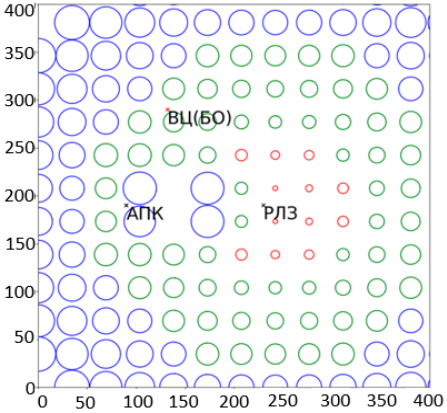


Рис. 2

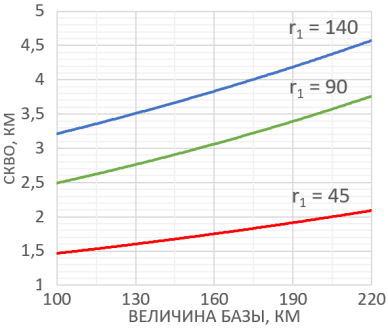


Рис. 3

3. Прием и декодирование некриптографических импульсных последовательностей ОС, передаваемых на РЛЗ в дискретно-адресном режиме работы MkXII (режиме S), позволяет извлекать в СРТР точные координаты ВЦ(БО), измеренные бортовой навигационной системой. и. В других режимах работы MkXII (режимы 1, 2, 3) дополнительно может извлекаться важная информация о высоте и скорости полета, государственной принадлежности ВЦ(БО), бортовом номере и др.

4. Важным достоинством АПМ является возможность определения плоскостных координат ВЦ (БО) при работе СРЛО в криптографическом режимах (4 и 5), так как для реализации метода знание ключа и структуры кодовых посылок сигналов СРЛО не требуется.

[1] Вопросы перспективной радиолокации. Коллективная монография / Под ред. Л.В. Соколова. – М.: Радиотехника, 2003, 512 с.
[2] Военное агентство стандартизации (MAS)1110. MAS/349-EL/4193 (Часть 1). – Брюссель, 1990, 93 с.
[3] «Ниобий-СВ» https://bigenc.ru/c/niobii-sv.
[4] Цикин И.А., Поклонская Е.С. Обработка сигналов системы вторичной радиолокации на удаленном пункте контроля // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2017. Т. 10, № 2. С. 58.
[5] Вторичный радиолокатор «Крона»: курс лекций / В.И. Коломиец, Н.П. Филимонов. – Красноярск.: Сибирский федеральный университет, 2007. 98 с.
[6] Дворников С.В., Саяпин В.Н., Симонов А.Н. Теоретические основы координатометрии источников радиоизлучений. Учебное пособие. – СПб.: ВАС, 2007, 80 с.
[7] Белоцерковский Г.Б. Основы радиолокации и радиолокационные устройства. – М.: «Сов. радио», 1975, 336 с.

Секция «Информационные системы.
Средства, технологии, безопасность»

Заседание секции проводилось 16 мая 2023 г.
Председатель – Л.Ю. Ротков, секретарь – А.А. Рябов.
Нижегородский государственный университет им. Н.И. Лобачевского.