ISSN 2074-7128(print), ISSN 2074-7136(on-line)

БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ
(IT Security)

Периодический рецензируемый научный журнал «Безопасность информационных технологий», освещающий широкий спектр проблем обеспечения информационной безопасности, в том числе технологические, организационно - правовые и образовательные аспекты.

Журнал зарегистрирован в Государственном комитете Российской Федерации по печати. Свидетельство №017789. Издается с 1994 г.

С момента основания и до настоящего времени учредителем журнала является федеральное государственное автономное образовательное учреждение высшего образования Национальный исследовательский ядерный университет «МИФИ» (НИЯУ МИФИ).

С 2007 г. и по настоящее время журнал входит в Перечень ВАК ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени доктора и кандидата наук по отраслям науки и группе специальностей научных работников 05.13.00 - информатика, вычислительная техника и управление, по которой журнал входит в этот перечень.

Основные тематические направления журнала:
•Концептуальные основы обеспечения информационной безопасности автоматизированных систем;
•Методические подходы к анализу и оценке рисков информационной безопасности, технологии поиска уязвимостей в программном обеспечении;
•Оценка уровня защищенности автоматизированных систем;
•Программно-технические способы и средства обеспечения информационной безопасности.

Журналом приветствуются статьи на русском и английском языках.

IT Security(Russia)

IT Security is a periodic peer-reviewed scientific journal publishing papers on a wide range of information security topics, including technological, organizational, legal and educational problems.

Since its establishment in 1994 (registration certificate No. 017789 by the State Committee for Press of the Russian Federation), the journal has been publishing by the Federal Autonomous Educational Institution of Higher Education National Research Nuclear University, a.k.a. "MEPhI" (Moscow Engineering Physics Institute).

Papers in Russian and English are equally welcome.

Focus topics:
•Fundamentals of information security of automated systems;
•Methodology of assessing the information security risks;
•Technology of detecting software vulnerabilities;
•Evaluation of the security level of automated systems;
•Soft - and hardware means of ensuring information security.

СОДЕРЖАНИЕ

Александр М. Алюшин, Сергей В. Дворянкин
ИСПОЛЬЗОВАНИЕ РЕЧЕВЫХ ТЕХНОЛОГИЙ ДЛЯ ЗАЩИТЫ ДОКУМЕНТООБОРОТА
6
Анастасия А. Баринова, Сергей В. Запечников
МЕТОДЫ И СРЕДСТВА ОБЕСПЕЧЕНИЯ КОНФИДЕНЦИАЛЬНОСТИ СМАРТ-КОНТРАКТОВ
16
Филипп Б. Буртыка
МЕТОДЫ ОРГАНИЗАЦИИ ПРОИЗВОЛЬНЫХ КОНФИДЕНЦИАЛЬНЫХ ЗАПРОСОВ К КРИПТОГРАФИЧЕСКОЙ ОБЛАЧНОЙ БАЗЕ ДАННЫХ С ПОМОЩЬЮ ПОЛНОСТЬЮ ГОМОМОРФНОГО ШИФРОВАНИЯ
24
Виктор С. Горбатов, Алексей А. Мещеряков
КУРС ТРЕНИНГА ПО БЕЗОПАСНОЙ РАЗРАБОТКЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
35
Александр М. Коротин
АНАЛИЗ УГРОЗ БЕЗОПАСНОСТИ ОТВЕТСТВЕННОЙ ИНФОРМАЦИИ, ПЕРЕДАВАЕМОЙ СИСТЕМОЙ АЛС НА БАЗЕ РАДИОКАНАЛА
42
Андрей Н. Кузнецов, Дмитрий А. Вышемирский
ОБ ОДНОМ ПОДХОДЕ К РЕШЕНИЮ ЗАДАЧИ ТОКЕНИЗАЦИИ ПРИ АНАЛИЗЕ БОЛЬШИХ МАССИВОВ ПОЛЬЗОВАТЕЛЬСКИХ ПАРОЛЕЙ
50
Владимир А. Минаев, Екатерина В. Вайц, Юлия В. Грачёва
ИМИТАЦИОННЫЕ ЭКСПЕРИМЕНТЫ С МОДЕЛЬЮ ИНФОРМАЦИОННО-ПСИХОЛОГИЧЕСКИХ ВОЗДЕЙСТВИЙ НА МАССОВОЕ СОЗНАНИЕ
61
Иван В. Нечта
АНАЛИЗ УСТОЙЧИВОСТИ ДИНАМИЧЕСКИХ ВОДЯНЫХ ЗНАКОВ
72
Алина В. Трепачева
О СООТНОШЕНИЯХ МЕЖДУ АТАКАМИ НА СИММЕТРИЧНЫЕ ШИФРЫ, ГОМОМОРФНЫЕ НАД КОЛЬЦОМ ВЫЧЕТОВ
82
АННОТАЦИИ
92
SUMMARY
98
СПИСОК СОКРАЩЕНИЙ, ПРИНЯТЫХ В ЖУРНАЛЕ
107

CONTENT

Alexandr M. Alyushin, Sergey V. Dvoryankin
THE USE OF SPEECH TECHNOLOGY TO PROTECT THE DOCUMENT TURNOVER
6

Anastasia A. Barinova, Sergey V. Zapechnikov
ON THE TECHNIQUES AND TOOLS FOR PRIVACY-PRESERVING SMART CONTRACTS
16

Filipp B. Burtyka
THE TECHNIQUES FOR ARBITRARY SECURE QUERING TO ENCRYPTED CLOUD DATABASE USING FULLY HOMOMORPHIC ENCRYPTION
24

Viktor S. Gorbatov, Alexey A. Meshcheriakov
SECURE SOFTWARE DEVELOPMENT TRAINING COURSE
35

Alexander M. Korotin
SECURITY THREAT ANALYSIS OF CRITICAL INFORMATION TRANSMITTED BY THE ALS SYSTEM BASED ON THE RADIO CHANNEL
42

Andrey N. Kuznetsov, Dmitry A. Vyshemirsky
ONE APPROACH TO SOLVING TOKENIZATION PROBLEM FOR ANALYSIS OF LARGE-SCALE COLLECTIONS OF USER-DEFINED PASSWORDS
50

Vladimir A. Minaev, Ekaterina V. Vaitc, Yulia V. Gracheva
SIMULATION EXPERIMENTS WITH THE MODEL OF INFORMATION-PSYCHOLOGICAL INFLUENCES ON MASS CONSCIOUSNESS
61

Ivan V. Nechta
ROBUSTNESS ANALYSIS OF DYNAMIC WATERMARKS
72

Alina V. Trepacheva
ON THE RELATIONS BETWEEN THE ATTACKS ON SYMMETRIC HOMOMORPHIC ENCRYPTION OVER THE RESIDUE RING
82

Александр М. Алюшин[1] , Сергей В. Дворянкин[2]
[1]Национальный исследовательский ядерный университет «МИФИ», Каширское ш., 31, г. Москва, 115409, Россия, e-mail: alyshin@list.ru
[2]Финансовый университет при Правительстве Российской Федерации, Ленинградский проспект, 49, Москва,125993, Россия, e-mail: SVDvoryankin@mephi.ru, ORCIDiD0000-0001-6908-0676

ИСПОЛЬЗОВАНИЕ РЕЧЕВЫХ ТЕХНОЛОГИЙ ДЛЯ ЗАЩИТЫ ДОКУМЕНТООБОРОТА.

DOI: http://dx.doi.org/10.26583/bit.2017.2.01

Аннотация. Несмотря на то, что мы живем в веке информационных технологий, бумажный документооборот не теряет своей актуальности. Выделены основные аспекты защиты документов, связанные с защитой их контекстной и юридической составляющих. К контекстной составляющей отнесен смысловой информационный аспект документа. К юридической – факты и условия создания, утверждения, согласования данного документа конкретными лицами. Показаны актуальность защиты документов в связи с возможными террористическими угрозами и важность для эффективной защиты документов такого фактора, как время выявления фальсификации. Представлен анализ требований к данному фактору для документов различного характера – финансовых, юридических, управляющих. Выделены управляющие документы, используемые для оперативного управления опасными объектами. Показано, что их умышленная фальсификация может привести к возникновению аварий и катастроф техногенного происхождения, а также человеческим жертвам. Дан сравнительный анализ применяющихся в настоящее время на практике методов защиты документов. Выделены биометрические и небиометрические методы защиты документов. Представлен анализ их недостатков. Сделан вывод о перспективности применения технологии защиты документов на основе речевой подписи. Рассмотрены основные этапы обработки голосовой информации при реализации данной технологии. Разработано программное обеспечение, реализующее новую технологию защиты документов от подделок. Сущность предлагаемой технологии заключается в том, что в конец документа добавляется аудиомаркер, хранящий в себе основную информацию защищаемого документа (финансовые аспекты, обязанности сторон, сроки и т.п.). Показано, что изменение текста в этом случае будет порождать необходимость изменения связанного с ним элемента защиты – аудиомаркера. Достоинством подхода является тот факт, что сделать это без автора или собственника документа невозможно, поскольку аудиомаркер хранит в себе биометрические данные указанного лица. Разработано приложение под мобильное устройство с операционной системой android, которое позволяет считывать, распознавать и воспроизводить аудиомаркеры. Проведенные испытания программного комплекса показали его высокую эффективность, в том числе при распознавании аудиомаркеров в условиях различной освещенности. Рассмотрена возможность повышения степени защищенности подлинности документа за счет фиксации в речевой подписи текущего уровня психоэмоционального напряжения автора документа, что может быть использовано для выявления случаев его неадекватного состояния. Это позволяет выявлять случаи составления документов в случае применения физического или психического насилия.

Ключевые слова: спектрограмма, аудиомаркер, распознавание, документация, графическая картинка.

Для цитирования. **АЛЮШИН, Александр М.; ДВОРЯНКИН, Сергей В.** Использование речевых технологий для защиты документооборота. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 6-15, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/100>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.01.

Благодарности. Данное исследование выполнено в рамках проекта «Апробация технологии снижения риска возникновения и уменьшение последствий катастроф техногенного происхождения за счет миниёмизации влияния человеческого фактора на надежность и безаварийность работы АЭС и других опасных объектов» в соответствии с госзаданием НИЯУ МИФИ на 2017-2019 гг.

Alexandr M. Alyushin[1] , Sergey V. Dvoryankin[2]
[1]National Nuclear Research University MEPHI (Moscow Engineering Physics Institute), Kashirskoeshosse, 31, Moscow, 115409, Russia, e-mail: alyshin@list.ru
[2]Finance University under the Government of Russia, Leningradsky Prospekt, 49, Moscow,125993, Russia, e-mail: SVDvoryankin@mephi.ru, ORCID iD0000-0001-6908-0676

The Use of Speech Technology to Protect the Document Turnover

DOI: http://dx.doi.org/10.26583/bit.2017.2.01

Abstract. The wide current paper documents implementation in practice workflows are shown. The basic aspects of document protection related to the protection of their content and legal components are underlined. For contextual component assigned semantic information aspect of the document is considered. For legal component attributed facts and conditions for the creation, approval, negotiation of the document to specific persons is viewed. The documents protection problem importance is shown in connection with possible terrorist threats. The importance of such factor as the time of fraud detection towards the efficiency of documents protection is shown. The fraud detection time requirements for documents of different nature – financial, legal, management is analyzed. The documents used for the operational management of dangerous objects is point out as the most sensitive to the falsification. It is shown that their deliberate falsification can lead to accidents and technogenic catastrophes and human casualties. A comparative analysis of currently used protecting documents methods are presented. Biometric and non-biometric methods of documents protection are point out. The analysis of their short comings are given. The conclusion about the prospects of document protection on the basis of the voice signature technology are done. The basic steps of voice information processing in the implementation of this technology are analyzed. The software that implements a documents counterfeiting new protection technology is proposed. The technology is based on the audiomarkers usage at the end of the document, which contains a general information about it. The technology is applicable to the wide range of documents such as financial and valuable papers, contracts, etc. One of the most important advantages of this technology is that any changes in the document can not be done without the author of the document because audiomarker keeps the biometric data of the person. The developed software is oriented on the mobile applications on the basis of android operating system. The software has a friendly interface, which allows one to create, correct and check audiomarkers. The performed software tests shows its high efficiency, including the possibility of audiomarkers recognition in various lighting conditions. The possibility of document protection level improvement by logging in voice signature the current level of mental and emotional stress of the document author, which can be used to detect its inadequate state is shown. This makes it possible to identify cases of document preparation under the application of physical or mental violence.

Keywords: spectrogram, audiomarker, recognition, documentation, graphic image

For citation. **ALYUSHIN, Alexandr M.; DVORYANKIN, Sergey V.** The Use of Speech Technology to Protect the Document Turnover. IT Security, [S.l.], v. 24, n. 2, p. 6-15, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/100>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.01.

Acknowledgements: This research was performed in the framework of the project «The Testing of technologies to reduce risk and reduction of consequences of catastrophes of technogenic origin due to minimization of human factor influence on reliability and trouble-free operation of nuclear power plants and other hazardous facilities" in accordance with the state assignment NRNU MEPhI for the 2017-2019.

Введение

Мы живем в веке высоких информационных технологий. Практически все сферы жизнедеятельности человека так либо иначе связаны с компьютерной техникой. Так, еще до рождения человека компьютерные технологии широко используются для медицинских диагностических целей, например, при ультразвуковых исследованиях. Все этапы обучения также непосредственно связаны, либо просто базируются на использовании компьютеров. В настоящее время практически вся медицинская диагностическая техника, весь финансовый документооборот, средства коммуникации базируются на хранении, обработке и передаче информации в цифровом виде.

Но, несмотря на это, бумажного документооборота меньше не становится. Действительно, заключение каких-либо контрактов, договоров или других соглашений все равно фиксируется на бумаге. Практически все финансовые операции также дублируются бумажными носителями. Аналогичным образом обстоит дело и с представлением результатов различного рода проверок, результатов испытаний, тестирований, аттестаций, расследований. А это означает, что вопросы, связанные с защитой документов (как в бумажном виде, так и цифровом) от различного рода фальсификаций и подделок, не потеряли своей актуальности.

Рассматривая проблему защиты документов, следует выделить следующие основные аспекты – защиту контекстной составляющей документов и защиту юридической составляющей. В первом случае речь идет о защите смысловой информационной составляющей документов. Во втором случае – о защите самого факта создания, утверждения, согласования данного документа конкретными лицами.

Особую актуальность вопросы защиты документов приобретают в связи с террористическими угрозами. Так, например, в случае осуществления физического захвата лиц, уполномоченных подписывать какие-либо документы юридического, финансового, технического, а также информационного характера, существующими средствами защиты практически невозможно своевременно обнаружить факты насильственного, то есть сфальсифицированного незаконного создания упомянутых выше документов.

Необходимо отметить, что одним из показателей эффективности применяемой технологии защиты документов является время выявления подделки. Требования к данному фактору могут принципиальным образом отличаться для различных документов. Так, требования к данному фактору для некоторых юридических документов, например, прав на недвижимую собственность, могут быть достаточно мягкими, так как собственность такого рода не может быть физически перемещена, либо спрятана быстро. По этой причине, как правило, имеется достаточное время на оспаривание сфабрикованных документов в суде без утраты самой собственности.

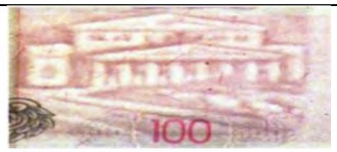
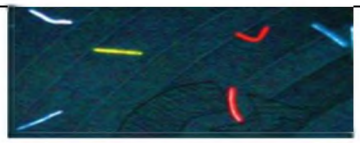
Иначе обстоит дело с финансовыми документами. Требования ко времени выявления факта фальсификации здесь достаточно жесткие. Это объясняется тем, что их фальсификация может привести к невосполнимой утрате финансов, так как денежные средства могут быть быстро переведены на подставные счета, с которых, даже в случае принятия необходимых решений суда, вернуть их будет просто невозможно.

Наиболее критична ситуация с документами, регламентирующими текущее управление опасными объектами. Показательными в этом плане являются, в первую очередь, тепловые и атомные станции, вредные производства, скоростной транспорт. Умышленная фальсификация документов такого рода может привести к авариям и катастрофам техногенного происхождения, человеческим жертвам, огромному экономическому урону. Для таких документов требования к эффективности защиты контекстной информации и юридических аспектов создания документа, а также времени

выявления фактов его подделки наиболее жесткие. Отсутствие фальсификации документов для рассмотренных выше объектов должно быть выявлено практически сразу же по их получению, либо перед исполнением содержащихся в них распоряжений.

Рассмотрим наиболее часто применяемые в настоящее время способы защиты бумажных документов [1, 2, 3]. Все способы защиты можно разделить на два раздела: биометрические и небиометрические методы защиты (таблица 1).

Таблица 1. − Небиометрические (слева) и биометрические (справа) методы защиты документов.

<table>
<tr><th colspan="2">Небиометрические методы защиты</th><th>Биометрические методы защиты</th></tr>
<tr><td>  </td><td>Защитные нити</td><td rowspan="3">Дактилоскопия
  </td></tr>
<tr><td>  </td><td>Водяные знаки</td></tr>
<tr><td>  </td><td>Защитные волокна</td></tr>
<tr><td>  </td><td>Цветопеременная краска</td><td>Подпись (почерк)
  </td></tr>
<tr><td>  </td><td>Орнаментальная полоса</td><td></td></tr>
</table>

Небиометрические методы защиты это: защитные нити, водяные знаки, защитные волокна, цветопеременные краски, орнаментальные полосы и др. Наглядный пример использования этих средств защиты можно увидеть на любой банковской купюре. Небиометрические методы защиты позволяют достаточно быстро выявить факт подделки документа.

Биометрические методы защиты документов − это отпечатки пальцев, подпись, почерк и др. Процесс выявления подделок документов в случае применения биометрических методов защиты требует несколько большего времени.

Из анализа биометрических и небиометрических методов защиты информации следуют два основных вывода. Первый − это то, что все они в той или иной мере могут быть подделаны (одни − более просто, другие − более сложно). Второй, и самый главный, −это то, что эти методы никак не связаны с информационным смысловым содержанием защищаемого документа. То есть, изменение сути договора не порождает изменение элементов защиты носителей этого договора.

Технология речевой подписи

Из вышеизложенного следует, что значительно повысить защищенность документов позволит использование таких биометрических технологий, которые будут связывать

защиту с самим защищаемым текстом. Одной из таких технологий является технология речевой подписи [4] (рисунок 1).

Данная технология заключается в следующем:

1. Автор защищаемого документа создает аудиофайл, в котором записан прочтенный им текст, содержащий основные моменты защищаемого документа (или весь документ целиком).
2. Автор загружает данный аудиофайл на специальный сайт для расчета и получения аудиомаркера (речевой подписи), выбирает типы границ (которые нужны для распознавания с помощью смартфона) и нажимает кнопку “Сгенерировать речевую подпись”.
3. Пользователь копирует изображение получившейся речевой подписи и вставляет в определенное место защищаемого документа, к примеру, в его конец.

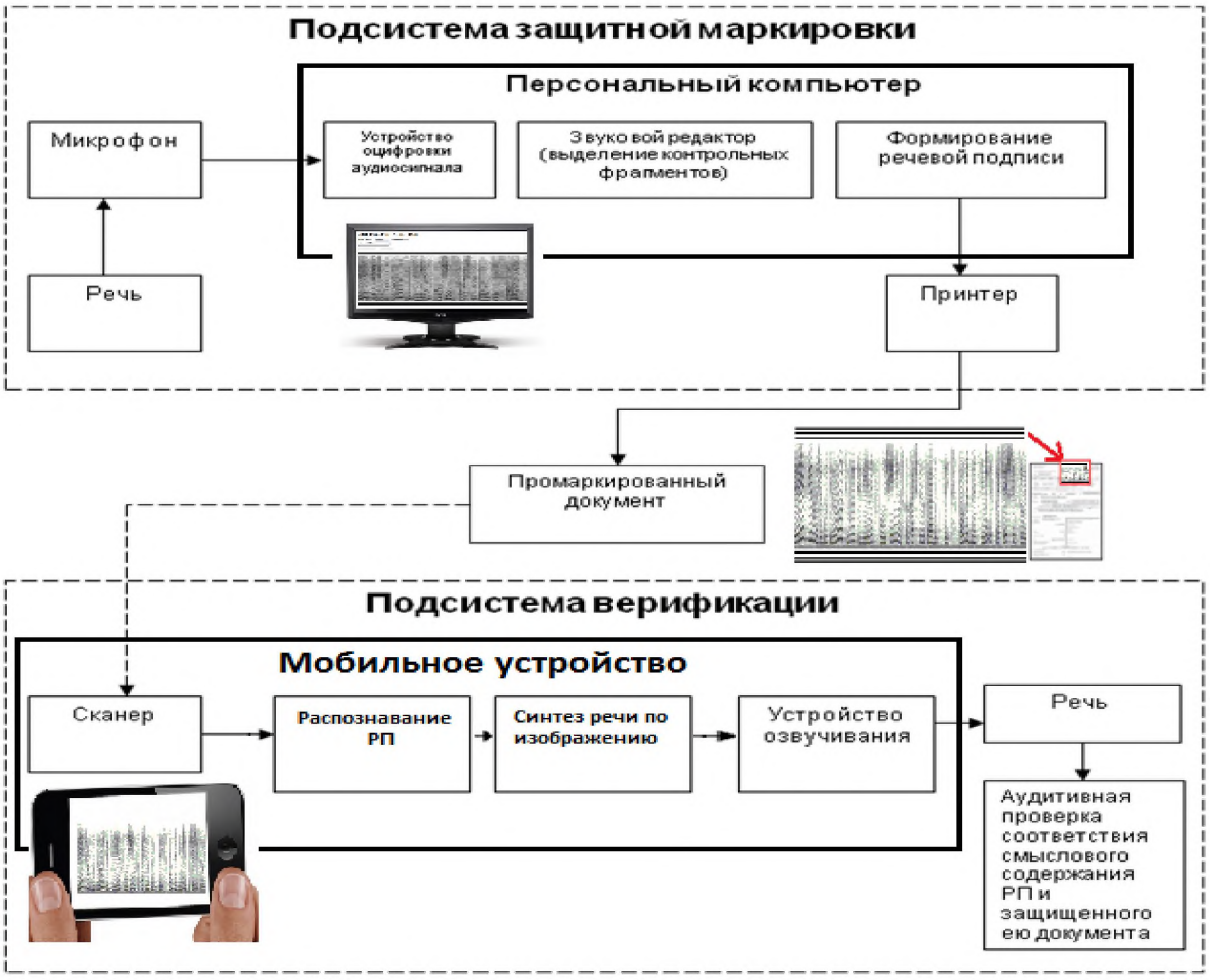


Рис.1. Технология речевой подписи
Fig.1. Voice signature technology

Далее пользователь скачивает и устанавливает специальное приложение на смартфон, с помощью которого можно распознать и воспроизвести изображение речевой подписи.

4. Пользователь запускает скаченное приложение, наводит камеру на речевую подпись (аудиомаркер), ждет несколько секунд, пока программа распознает речевую подпись и просинтезирует ее, и после этих действий пользователь может прослушать получившийся результат и дать заключение о совпадении или несовпадении

аудиоинформации с текстом документа, а также голоса человека, который заключал договор, с голосом человека, который представлен в речевой подписи.

Иными словами, при использовании технологии речевой подписи (РП), можно утверждать, что злоумышленник не сможет изменить содержимое документа без изменения изображения речевой подписи (аудиомаркера), в которой хранятся основные моменты смыслового содержания защищаемого документа и биометрические данные его собственника или автора.

Рассмотрим более подробно, что же такое речевая подпись. По сути, этот аудиомаркер представляет собой изображение спектрограммы акустического (речевого) сигнала с границами для распознавания (таблица 2).

Для получения спектрограммы звукового сигнала используется математический аппарат кратковременного анализа Фурье (КАФ). В основу кратковременного анализа положено допущение, что в пределах относительно короткого интервала речь можно считать стационарным процессом. Это позволяет отражать в общем виде описания сигнала, его важнейшие временные изменения [4, 5, 6].

Спектральная плотность сигнала при построении РП вычисляется по формуле (1):

$$S(\omega,t)=\sum_{n=\frac{N-M}{2}}^{\frac{N+M}{2}-1} w(n)x(n+st\cdot k)\,e^{-\frac{2\pi i}{N}t\cdot n}, \qquad (1)$$

где $\omega = 2\pi i\frac{\nu}{N}$; $t = \frac{st\cdot k}{\nu}$;
$\omega(n)$ − оконная функция;
$x(n)$ −дискретизированный аудиосигнал;
$S(\omega,t)$ − спектральная плотность сигнала на частоте ω в момент времени t;
$\nu = 8$ кГц − частота дискретизации аудиосигнала;
st − шаг Фурье-анализа;
N=1024 − база Фурье;
M=512 − ширина окна.

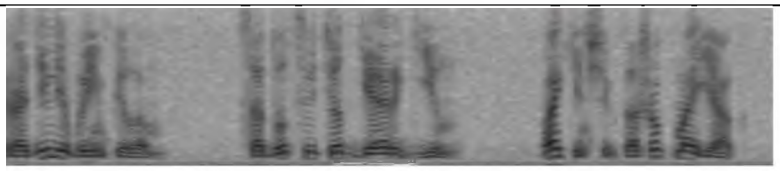
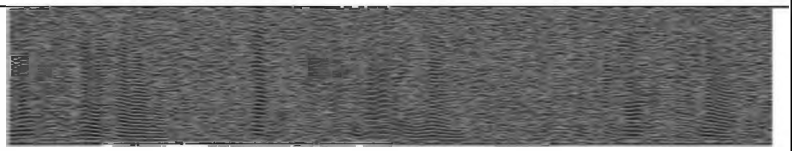
Для анализа речевого сигнала его Фурье-образ является удобным по двум причинам [6]. Во-первых, через представление РС в виде синусоид или комплексных экспонент очень удобно определять их отклик при прохождении сигнала через линейные системы с помощью основных свойств Фурье-преобразования. Во-вторых, на Фурье-образе часто выявляются такие свойства исходного сигнала, которые в первоначальном виде скрыты или не очевидны. В качестве оконной функции было выбрано окно Гаусса (2) с параметром $\sigma = 0.17$, так как оно давало лучший результат по сравнению с другими окнами.

$$w(n)=e^{-\frac{1}{2}\left(\frac{n-A}{A\sigma}\right)^2}, \text{ где } A=\frac{M-1}{2} \qquad (2)$$

В таблице 2 представлены осциллограмма, спектрограмма и фазограмма участка речевого сигнала, наблюдаемого в один и тот же момент времени. Осциллограмма – это зависимость амплитуды сигнала от времени, спектрограмма показывает зависимость интенсивности сигнала от частоты в конкретный момент времени (3), а фазограмма − зависимость косинуса фазы от частоты и времени. Аудиомаркер РП в виде изображения спектрограммы можно получить, рассчитав спектральную плотность сигнала [5,7, 8, 9]:

$$p[i][j] = 255 * \left|\sin(\arg(S[i][j]) - \arg(S[i][j-1])) - s * \frac{2\pi}{N} i\right| \quad (3)$$

Таблица 2 – Изображения осциллограммы, спектрограммы, фазограммы речевого сигнала

Осциллограмма	
Спектрограмма (сонограмма – для человеческой речи)	
Фазограмма	

Соответственно, если к такой спектрограмме привязать границы для распознавания и озвучивания аудиомаркера при помощи смартфона, то в результате мы получим изображение речевой подписи [9] (рисунок 2).

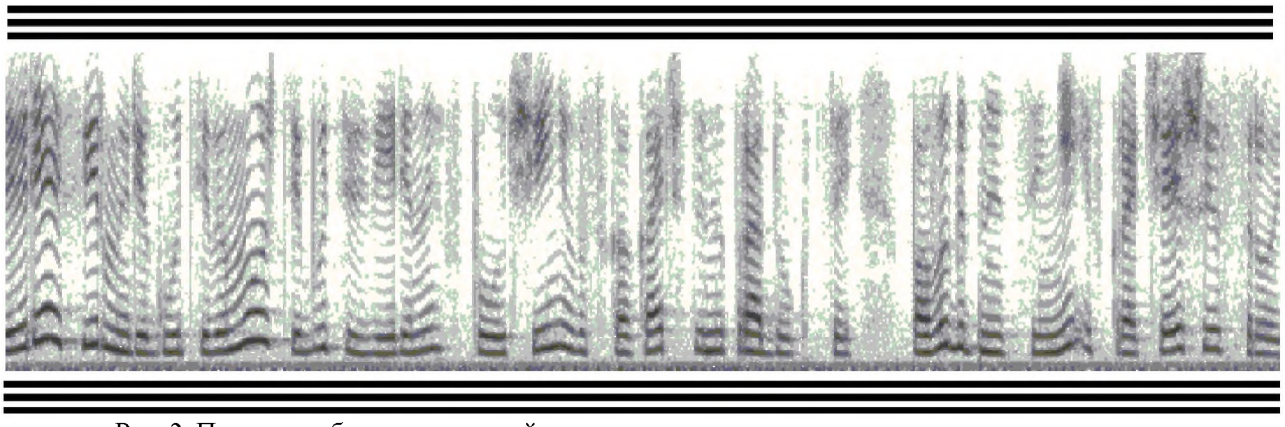


Рис. 2. Пример изображения речевой подписи как аудиомаркера защищаемого документа
Fig. 2. An example of voice signature image as audiomarker of protected document

Анализ технологии выбора границ и алгоритм распознавания речевой подписи подробно представлен в статьях [9, 10].

Применение рассматриваемой технологии речевой подписи дает возможность осуществить идентификацию автора документа на основе анализа спектра его голоса [11-15]. Для этого необходимо иметь образцы голоса в базе данных [16]. Кроме этого, разработанные к настоящему времени акустические технологии определения текущего психоэмоционального состояния говорящего [17-20], позволяют выявить факты неадекватного состояния человека, составившего и подписавшего документ, например, в

состоянии алкогольного, либо наркотического опьянения, в случае физического и психологического принуждения.

Таким образом, рассматриваемая технология речевой подписи позволяет защитить как контекстную составляющую документа, так и юридические аспекты условий его создания. Последний фактор можно рассматривать как дополнительную степень защиты документа.

Данная технология также удовлетворяет жестким требованиям ко времени выявления факта подделки документа и может быть рекомендована к защите практически всех типов документов – юридических, финансовых, управляющих.

Заключение

Бумажный документооборот остается в настоящее время одним из наиболее распространенных способов фиксации, хранения и передачи правовой, финансовой, политической, научно-технической и других видов информации.

Проведенный анализ существующих биометрических и небиометрических методов идентификации для защиты документированной информации позволил выявить области возможного их применения, а также наиболее существенные недостатки, ограничивающие их эффективность, в первую очередь, по причине возможной фальсификации, а также независимости технологий защиты от содержания документа.

Предложена технология защиты документов на основе использования аудиомаркера в виде изображения спектрограммы (речевой подписи) как нового средства защиты документооборота, имеющего существенные преимущества по сравнению с существующими биометрическими методами.

Показана возможность использования аудиомаркера для решения задачи идентификации говорящего, а также выявления случаев его неадекватного поведения, вызванного, в том числе, применением физического и психического насилия, либо состоянием алкогольного, либо наркотического опьянения.

Разработаны алгоритмы построения и распознавания речевой подписи для мобильного устройства. Создано специализированное программное обеспечение, ориентированное на работу под операционной системой android.

Проведенное лабораторное тестирование разработанных методических и технических средств подтвердило высокую эффективность технологии речевой подписи. Рассмотренная технология может быть использована для повышения степени защиты широкого спектра документов финансового, юридического, экономического и технического характера и гарантирует минимальное время выявления фактов подделки и фальсификации данных документов.

СПИСОК ЛИТЕРАТУРЫ:

1. Шашкин С.Б. Технико-криминалистические исследования документов со специальными средствами защиты от подделки: Учебное пособие. Саратов: СЮИ МВД России. 2002. − 96 с.
2. Гудилин Д. Печатные технологии и защита документов от подделки. Компьюарт, №11, 2003 2. 2000 - 2009.
3. Коншин А.А. Защита полиграфической продукции от фальсификации. М.: ООО «Синус», 2000.
4. ANSI S3.5-1997, American National Standard Methods for Calculation of the Speech Intelligibility Index – American National Standards Institute, New York. – 1997.
5. Дворянкин С.В. Речевая подпись. М. РИО. 2003. – 184 с.
6. Рабинер Л.Р. Цифровая обработка речевых сигналов/ Л.Р. Рабинер, Р.В. Шафер: Пер. с англ. / Под ред. М.В. Назарова и Ю.Н. Прохорова. М.: Радио и связь.– 1981. – 496 с.
7. Ephraim Y.A. Brief Survey of Speech Enhancement/ Y. Ephraim, H. Lev-Ari, W.J.J. Roberts // The Electronic Handbook, CRC Press. – 2005.
8. Vorobiov V.I. Inter Component Phase Processing of Speech Signals for Their Recognition and Identification of Announcers/ V.I. Vorobiov // XVIII Session of the Russian Acoustical Society, Taganrog, September 11-15, 2006.

9. Алюшин А.М., Дворянкин Н.С. Особенности распознавания изображений речевой подписи на мобильных устройствах //Безопасность информационных технологий. 2015. № 4. С. 38-45.
10. Алюшин А.М., Дворянкин Н.С. Технология защитного аудиомаркирования документированной информации с использованием мобильных устройств // Спецтехника и связь. 2015. № 6. с. 26-31.
11. Bimbot F. et al. A Tutorial on Text-Independent Speaker Verification. - EURASIP Journal on Applied Signal Processing, 2004, No4. P. 430–451.
12. Reynolds D. Experimental evaluation of features for robust speaker identification. – IEEE Trans. On Speech and Audio Processing, 1994, vol. 2, No4. P. 639-643.
13. Коваль С.Л., Лабутин П.В., Малая Е.В., Прощина Е.А. Идентификация дикторов на основе сравнения статистик основного тона голоса //В сб.трудов XV международной Научной конференции «Информатизация и информационная безопасность правоохранительных органов». –М.: Академия управления МВД России , 2006. С. 324-327.
14. Vogt R., Baker B., Sridharan S. Modeling session variability in text-independent speaker verification // In Proc. Eurospeech, Lisbon, Portugal, Sept. 2005, pp. 3109–3112.
15. Rosenberg A. et al. Cepstral channel normalization techniques for HMM-based speaker verification // In Proc. ICSLP-94, pp. 1835-1838.
16. Тимофеев А.В. Распределённая система фоноучёта «VoiceNet ID». – Речевые технологии , 2009, No 2, с. 69-73.
17. Алюшин В.М. Спектральный анализ речевой деятельности как способ оценки психологического климата в коллективе // Вопросы психологии, 2016. № 3. С.148–156.
18. Алюшин В.М. Диагностика психоэмоционального состояния на основе современных акустических технологий // Вопросы психологии, 2015. № 3. С.145–152.
19. Алюшин М.В., Колобашкина Л.В. Мониторинг биопараметров человека на основе дистанционных технологий // Вопросы психологии, 2014. № 6. С.135–144.
20. Алюшин М.В. и др. Акустические технологии для интеллектуальных систем мониторинга функционального состояния оперативного состава управления объектами атомной энергетики /М.В. Алюшин, В.М. Алюшин, С.В. Дворянкин, Л.В. Колобашкина //Глобальная ядерная безопасность. 2013. № 4(9). С. 63-71.

REFERENCES:

[1] Shashkin S.B. Technical-criminalistics research of documents with special means of protection against forgery: a tutorial. Saratov: SYUI MVD Russia, 2002. − 96 p. (in Russian).
[2] Gudilin D. Printing technologies to protect documents against forgery. Komp'yuart, №11, 2003 2. 2000 - 2009. (in Russian).
[3] Konshin A.A. Protection of printed products against falsification. M.: OOO «Sinus», 2000. (in Russian).
[4] ANSI S3.5-1997, American National Standard Methods for Calculation of the Speech Intelligibility Index – American National Standards Institute, New York. – 1997.
[5] Dvoryankin S.V. Voice signature. M. RIO. 2003, 184 p. (in Russian).
[6] Rabiner L.R. Digital processing of speech signals; L.R. Rabiner, R.V. Shafer: transl. from English. Pod red. M.V. Nazarova i Y.N. Prohorova. M.: Radio isvyaz'.– 1981. – 496 p.
[7] Ephraim Y.A. Brief Survey of Speech Enhancement; Y. Ephraim. H. Lev-Ari, W.J.J. Roberts; The Electronic Handbook, CRC Press. – 2005.
[8] Vorobiov V.I. Inter Component Phase Processing of Speech Signals for Their Recognition and Identification of Announcers; V.I. Vorobiov. XVIII Session of the Russian Acoustical Society, Taganrog, September 11-15, 2006.
[9] Alyushin A.M., Dvoryankin N.S. Features image recognition of voice signatures on mobile devices. IT Security (Russia). v.22, 4(2015), P. 38-45. (in Russian).
[10] Alyushin A.M., Dvoryankin N.S. Technology protective audiomaterialy documented information using mobile devices. Spectekhnika i svyaz'. 2015. № 6. P. 26-31. (in Russian).
[11] Bimbot F. et al. A Tutorial on Text-Independent Speaker Verification. - EURASIP Journal on Applied Signal Processing, 2004, No4. P. 430–451.
[12] Reynolds D. Experimental evaluation of features for robust speaker identification. – IEEE Trans. On Speech and Audio Processing, 1994, vol. 2, No4. P. 639-643.
[13] Koval' S.L., Labutin P.V., Malaja E.V., Proshhina E.A. Speaker identification based on the comparison of statistics primary tone of voice. V sb. trudov XV mezhdunarodnoj Nauchnoj konferencii «Informatizacija i informacionnaja bezopasnost' pravoohranitel'nyh organov» – M.: Akademija upravlenija MVD Rossii , 2006. P. 324-327.(in Russian).
[14] Vogt R., Baker B., Sridharan S. Modeling session variability in text-independent speaker verification; In Proc. Eurospeech, Lisbon, Portugal, Sept. 2005. P. 3109–3112.
[15] Rosenberg A. et al. Cepstral channel normalization techniques for HMM-based speaker verification; In Proc. ICSLP-94. P. 1835-1838.

[16] Timofeev A.V. Distributed system of Pinocheta «VoiceNet ID». – Rechevye tehnologii, 2009, No 2. P. 69-73. (in Russian).
[17] Alyushin V.M. Spectral analysis of speech as a means of assessing psychological team climate. Voprosy Psikhologii. 2016. N 3. P.148–156. (in Russian).
[18] Alyushin V.M. Diagnostics of emotional states on the basis of contemporary acoustic technologies. Voprosy Psikhologii. 2015. N 3. P. 145–152. (in Russian).
[19] Alyushin M.V., Kolobashkina L.V. Monitoring human biometric parameters on the basis of distance technologies. Voprosy Psikhologii. 2014. N 6. P.135–144. (in Russian)
[20] Alyushin M.V. i dr. Acoustic technologies for "intellectual" monitoring systems of atomic energetic objects' operational control staff current functional state./M.V. Alyushin, V.M. Alyushin, S.V. Dvoryankin, L.V. Kolobashkina. Global'naya yadernaya bezopasnost'. 2013. N 4(9). P. 63-71. (in Russian).

Поступила в редакцию – 19 февраля 2017 г. Окончательный вариант – 20 мая 2017 г.
Received – February 19, 2017. The final version – May 20, 2017.

Анастасия А. Баринова, Сергей В. Запечников
Национальный исследовательский ядерный университет «МИФИ», Каширское шоссе, 31, г. Москва, 115409, Россия,
e-mail: naicqa@yandex.ru, ORCID iD 0000-0002-3784-6224,
e-mail: SVZapechnikov@mephi.ru , ORCID iD 0000-0002-7975-6040

МЕТОДЫ И СРЕДСТВА ОБЕСПЕЧЕНИЯ КОНФИДЕНЦИАЛЬНОСТИ СМАРТ-КОНТРАКТОВ

DOI: http://dx.doi.org/10.26583/bit.2017.2.02

Аннотация. В настоящее время наблюдается тенденция качественного и количественного усложнения бизнес-процессов. Электронная форма обмена информацией становится преобладающим средством коммуникаций, а также одним из основных элементов обеспечивающей инфраструктуры для бизнес-процессов. Эти обстоятельства обеспечивают необходимость внедрения электронных смарт-контрактов («умных контрактов») – протоколов, которые описывают набор условий и автоматически следят за их выполнением. Наиболее важным требованием к подобным технологиям является обеспечение конфиденциальности данных. В данной работе были исследованы наиболее известные платформы конфиденциальных электронных контрактов, выявлены методы обеспечения конфиденциальности, критерии их сравнения, проведено сравнение по выделенным критериям и сформулированы выводы о требованиях к платформам электронных контрактов, обеспечивающих конфиденциальность.

Ключевые слова: смарт контракты, блокчейн, конфиденциальность

Для цитирования. **БАРИНОВА, Анастасия А.; ЗАПЕЧНИКОВ, Сергей В.** Методы и средства обеспечения конфиденциальности смарт-контрактов. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 16-23, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/101>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.02..

Anastasia A. Barinova, Sergey V. Zapechnikov
National Research Nuclear University MEPHI, Kashirskoe sh., 31, Moscow, 115409, Russia,
e-mail: naicqa@yandex.ru, ORCID iD 0000-0002-3784-6224,
e-mail: SVZapechnikov@mephi.ru , ORCID iD 0000-0002-7975-6040

On the techniques and tools for privacy-preserving smart contracts

DOI: http://dx.doi.org/10.26583/bit.2017.2.02

Abstract. Currently business processes become more and more complicated. Data used in these processes circulates mainly through the digital communications. Due to these conditions some kind of electronic contracts for business deals become necessary. Smart contracts should describe a set of conditions, implemented through some events in the real world and digital systems. The most important requirement for this technology is privacy ensuring. In this work we have explored existing projects of privacy-preserving smart contracts defined comparison criteria compared projects and made a conclusion about options required for smart contract frameworks.

Keywords: smart contracts, privacy, blockchain

For citation. **BARINOVA, Anastasia A.; ZAPECHNIKOV, Sergey V.** On the Techniques and Tools for Privacy-Preserving Smart Contracts. IT Security (Russia), [S.l.], v. 24, n. 2, p. 16-23, June 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/101>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.02.

Введение

Усложнение бизнес-процессов в современном мире влечет за собой необходимость формирования четкого и справедливого аппарата их регулирования, который обеспечит

отсутствие возможности вмешательства третьих лиц в выполнение контракта и снизит риски неоднозначных трактовок условий контракта [1].

Идея внедрения подобной технологии, получившей название смарт-контракт («умный контракт», интеллектуальный контракт), была впервые предложена американским программистом Ником Сабо в 1994 году [2]. Однако для ее практической реализации не существовало необходимых технологических средств.

В связи с активным развитием технологий децентрализованных сетей стало возможным создание платформ для реализации и выполнения смарт-контрактов [3]. Крупные научно-исследовательские центры, университеты и отдельные команды энтузиастов разработали целый ряд платформ смарт-контрактов: Ethereum, Counterparty, Enigma и др. Однако они не решали одну из основных проблем смарт-контрактов: учитывая тот факт, что при заключении и выполнении контрактов стороны зачастую предоставляют множество персональных данных и (или) данных, составляющих коммерческую тайну, одним из важнейших требований к платформам смарт-контрактов является обеспечение конфиденциальности данных. Это требование обуславливает актуальность проблемы обеспечения конфиденциальности смарт-контрактов. В то же время все ранее известные платформы смарт-контрактов основывались на технологии распределенного нередактируемого реестра – блокчейн, который является открытым и доступен всем членам сети. Следовательно, ни о какой конфиденциальности данных говорить в таком случае не приходится.

Технологии распределенного нередактируемого реестра и смарт-контрактов

Для рассмотрения вопросов обеспечения конфиденциальности смарт-контрактов, необходимо выявить основные особенности тех технологий, на которых они основываются, – технологии распределенного нередактируемого реестра, или блокчейна, и классической технологии смарт-контрактов.

Технология распределенного нередактируемого реестра позволяет пользователям взаимодействовать друг с другом в децентрализованной среде без необходимости взаимного доверия и участия третьих лиц [3]. Каждый узел сети обладает специальным реестром («ledger»), который представляет собой полную копию с записями всех транзакций, проводимых участниками сети. Записи в реестр можно вносить лишь с согласия большинства пользователей, а однажды записанная информация не может быть удалена или модифицирована [4]. В реестре данные группируются в блоки, которые, в свою очередь, формируют цепочки блоков – отсюда и название блокчейн (blockchain) – буквально «цепочка блоков». Блоки в распределенном реестре связаны с помощью хэш-кодов, поскольку в заголовке каждого последующего блока хранится хэш-код предыдущего блока. Именно это техническое решение обеспечивает защищенность и устойчивость реестра к несанкционированному редактированию [5].

Смарт-контракт – это часть прикладной программы, которая хранит правила согласования условий контракта, автоматически следит за исполнением контракта и выполняет установленные условия контракта [6].

Основная идея смарт-контракта заключается в том, что договорное управление той или иной сделкой между двумя или более сторонами может осуществляться автоматически (программным способом) через распределенный реестр, и для этого не требуется привлечение какой-либо доверенной третьей стороны (ДТС): арбитра либо третейского судьи [7].

Таким образом решаются проблемы традиционного договорного права — зависимость сторон контракта от ДТС, а также временные и материальные затраты на коммуникации между сторонами контракта и ДТС.

Смарт-контракт реализует получение и отправку любых данных пользователей (“value”) с помощью транзакций, а также изменение своего внутреннего состояния (“state”) посредством событий (рис. 1).

В идеале смарт-контракт должен выполняться в среде, которая позволяет полностью автоматизировать процесс его выполнения, то есть в среде, имеющей беспрепятственный доступ исполняемого кода к объектам договоренностей, фигурирующим в тексте (условиях) смарт-контракта [8]. Следовательно, все условия такого смарт-контракта должны иметь четкое логико-математическое описание.

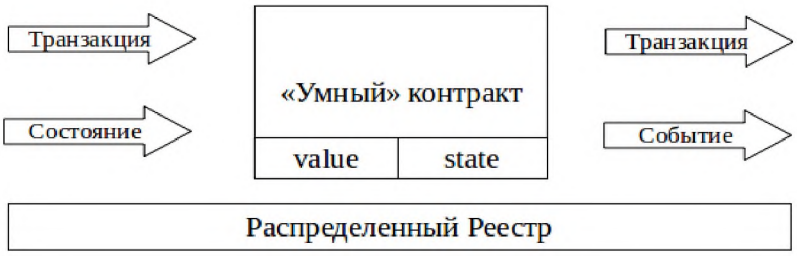


Рис. 1. Общая схема функционирования смарт-контракта
Fig. 1. The General scheme of the functioning of the smart-contract

Платформы конфиденциальных смарт-контрактов

Благодаря многочисленным исследованиям в области обеспечения конфиденциальности при заключении смарт-контрактов многими университетами и научными центрами за последние годы были запущены первые проекты, реализующие эту технологию.

Рассмотрим основные идеи двух наиболее интересных проектов — платформы Hawk, созданной на базе Мэрилендского и Корнельского университетов (США), а также платформы Enigma, реализованной исследователями из Массачусетского технологического института (США).

Платформа Hawk. Основная идея, положенная в основу платформы конфиденциальных смарт-контрактов Hawk, заключается в трансляции с помощью собственного компилятора обычной программы смарт-контракта в криптографический протокол взаимодействия пользователей с распределенным реестром [9].

Откомпилированная программа Hawk состоит из двух частей: открытой и закрытой (рис. 2). Закрытая часть непосредственно взаимодействует с данными пользователей, а также проводит вычисления для определения выплат между сторонами контракта. Ее главная функция — обеспечить защищенность данных пользователей, а также денежных потоков. Открытая программная часть не осуществляет взаимодействие с данными пользователей.

Выполнение программы смарт-контракта контролируется специальной стороной, называемой менеджером, который имеет доступ к пользовательским входным данным и обязан не разглашать их [9]. Однако очень важно отметить, что менеджер не соотносится с понятием ДТС — даже в том случае, если менеджер произвольно отклоняется от выполнения протокола либо вступает в сговор с одной из сторон, он не может повлиять на корректность исполнения контракта. В случае прерывания контракта менеджером он будет финансово наказан [10]. Каждая программа на платформе Hawk при работе использует показания специального таймера, который определяет время и очередность наступления событий.

Конфиденциальность контрактов на платформе Hawk обеспечивается за счет следующих особенностей.

1. Обеспечивается конфиденциальность данных каждого отдельно взятого контракта по отношению к внешней среде. Несмотря на то, что стороны контракта обмениваются данными с реестром, денежные потоки и данные транзакций из закрытой части программы контракта защищены от внешней среды криптографическими методами. В распределенный реестр направляется зашифрованная информация, а доказательство с нулевым разглашением [11] обеспечивает соблюдение корректности выполнения контракта и взаимодействия с данными [10].

2. Обеспечивается конфиденциальность данных внутри одного контракта. Платформа Hawk предполагает, что стороны контракта защищают свои собственные финансовые интересы. В частности, они могут произвольно отклоняться от определенного протокола либо преждевременно прерывать контракт. При этом Hawk обеспечивает не только конфиденциальность и аутентичность данных пользователей, но и финансовую справедливость в случае прерывания сделок. Это достигается с помощью специального механизма возврата денежных средств после достижения определенных временных меток [10].

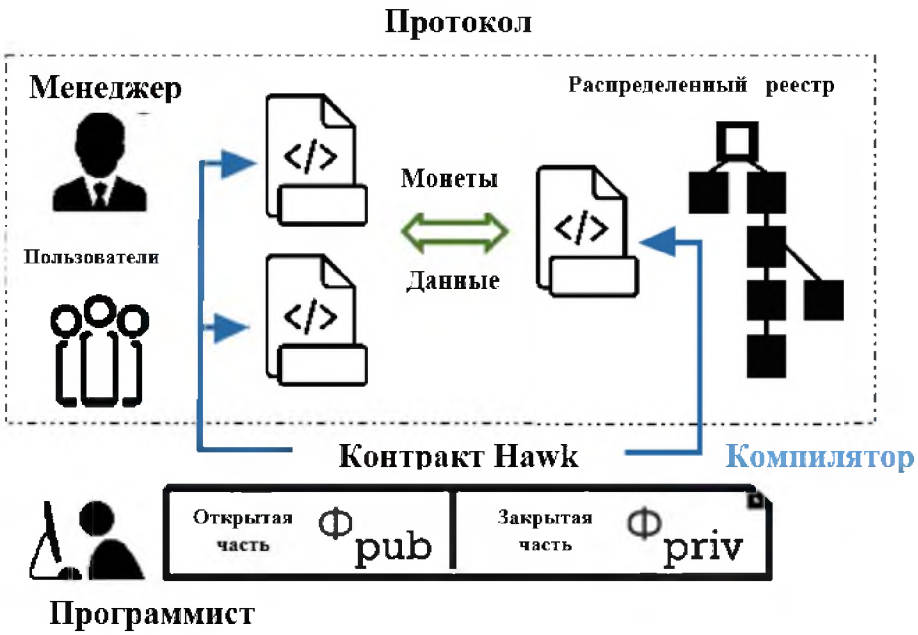


Рис. 2. Схема работы программы Hawk (по материалам [10])
Fig. 2. A scheme of the Hawk program functioning [10]

Платформа Enigma. Эта платформа позиционируется разработчиками как распределенная вычислительная платформа с гарантированным обеспечением конфиденциальности [12]. Основные свойства, которые обеспечивает платформа, заключаются в следующем.

1. *Конфиденциальность*. С помощью безопасных распределенных вычислений (sMPC — secure multi-party computation) работа с данными на этой платформе осуществляется вообще без участия ДТС. Данные разделяются между узлами сети, и те выполняют операции лишь со своей частью данных, которая представляет собою лишь бессмысленный фрагмент [12].

2. *Масштабируемость*. В отличие от традиционной формы распределенного реестра, вычисления, необходимые при выполнении контракта, не дублируются в каждом узле сети. Точно так же не дублируются многократно и данные, участвующие в выполнении контракта. Кроме того, это позволяет платформе Enigma выполнять вычисления над зашифрованными данными контракта без доступа к открытому тексту.

Интерпретатор разбивает процесс выполнения конфиденциального контракта, как показано на рис. 3, позволяя уменьшить время выполнения и при этом сохранив конфиденциальность [13].

Собственное хранилище данных Enigma осуществляет взаимодействие с распределенным реестром посредством использования схем разделения секрета и распределенных вычислений (рис. 4). Для этого вне блокчейна используется распределенная хэш-таблица (Distributed Hash Table – DHT) [14], которая доступна через реестр. При этом в реестре хранятся уже не сами данные, а ссылки на них. Личные данные должны быть зашифрованы на клиентской стороне перед началом взаимодействия с хранилищем и выполнением протоколов доступа [13].

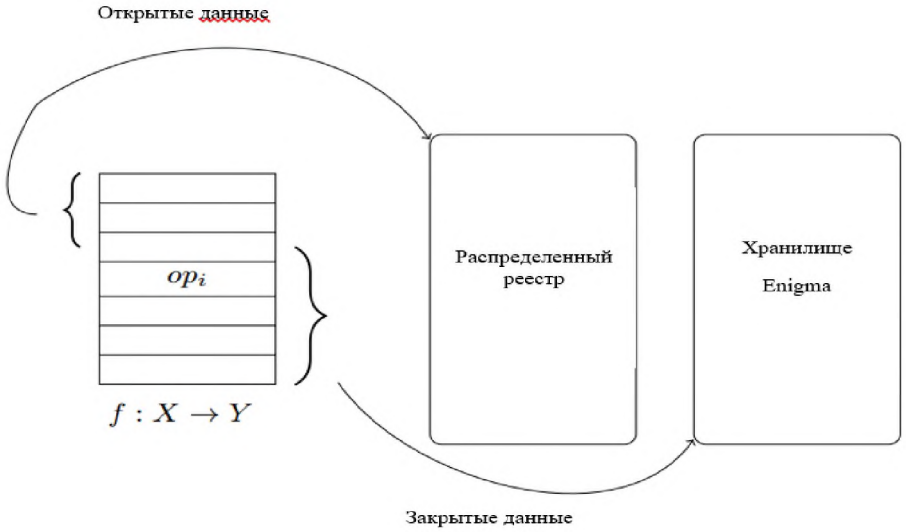


Рис. 3. Схема функционирования платформы распределенного реестра Enigma
Fig. 3. The scheme of functioning of the distributed registry platform Enigma

На сетевом уровне хранилище работает по протоколу Kademlia DHT [15] с использованием широковещательных каналов связи и схем шифрования с открытым ключом.

Сеть, построенная на базе Enigma, может выполнять код без утечки исходных данных к какому-либо узлу сети с помощью схемы линейного разделения секрета [13]. Для безопасных распределенных вычислений необходимо, чтобы каждый узел сети взаимодействовал с другим с коммуникационной сложностью$O(n^2)$ и с постоянным числом раундов. В случае линейных схем разделения секрета [16] эта вычислительная сложность в основном обусловлена операциями умножения, в то время как операции сложения могут выполняться параллельно, без обмена данными [12].

Сравнение платформ конфиденциальных смарт-контрактов

Сравнение двух платформ конфиденциальных смарт-контрактов – Hawk и Enigma – позволяет сделать следующие основные выводы.

1. Используемый платформой Hawk механизм доказательств с нулевым разглашением является более пригодным для использования в качестве средства обеспечения конфиденциальности по сравнению со схемой линейного разделения секрета в сочетании с безопасными распределенными вычислениями на платформе Enigma, так как доказательство с нулевым разглашением является широко апробированным и легко реализуемым криптографическим примитивом.

2. Принципиальная схема создания и выполнения контракта в реализациях Hawk и Enigma совпадает: контракт программируется, создается криптографические и иные протоколы его выполнения и взаимодействия пользователей, после чего он выполняется.

Однако на каждом этапе этого процесса различаются используемые методы. В схеме Hawk контракт после создания выполняется в распределенном реестре, в то время как в схеме Enigma контракт выполняется распределенным образом и в двух местах: распределенном реестре и хранилище Enigma. На стадии выполнения контракта Hawk большую роль играет специальный участник сети — менеджер, который обеспечивает корректное выполнение контракта, проводит основные операции с криптовалютой, отправляет данные пользователей в распределенный реестр и получает данные из него. В платформе Enigma выполнение контракта и гарантия его корректного исполнения обеспечивается за счет распределенных вычислений и специального протокола , который сопровождает такие вычисления.

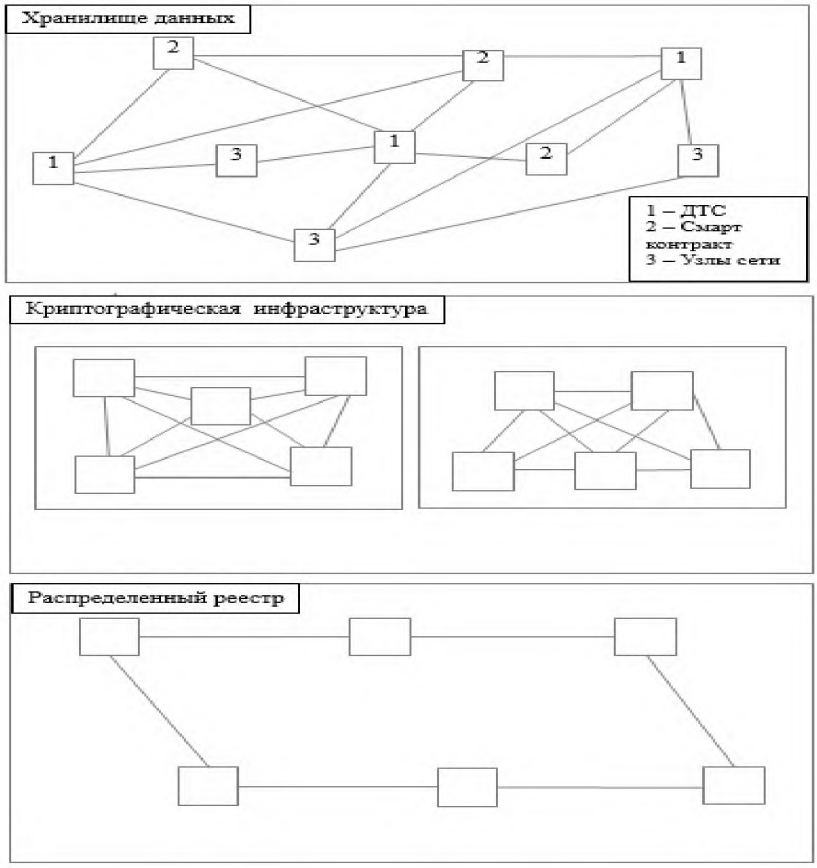


Рис. 4. Архитектура платформы Enigma (по материалам [17])
Fig. 4. The architecture of the platform Enigma [17]

3. В отличие от схемы хранения данных в платформе Hawk, схемы Enigma не предполагает стандартного хранения данных – дублирования данных реестра у каждого участника сети. Данный факт позволяет повысить производительность работы платформы Enigma за счет схемы доступа к данным через распределенный реестр и распределенную хэш-таблицу (DHT).

4. В отличие от платформы Enigma, контракты на платформе Hawk могут программироваться и сопровождаться людьми, не имеющими знаний в сфере криптографии и программирования.

5. Финансовая справедливость достигается в обеих схемах. Платформа Hawk использует для этого специальные временные метки и таймеры, которые определяют события в системе и контролируют наступление новых. Платформа Enigma реализует эти характеристики за счет распределенных вычислений.

Заключение

В результате данной работы:

- проведен обзор на наиболее известные платформы смарт-контрактов, обеспечивающих конфиденциальность, Hawk и Enigma;
- исследованы возможности обеспечения конфиденциальности смарт-контрактов в проектах Enigma и Hawk и их детальный анализ;
- выявлены основные методы реализации распределенного реестра, смарт-контракта;
- проведено сравнение платформ по выделенным критериям;

На основании сравнения сформулированы выводы, позволяющие выделить критерии, необходимые для проектирования и разработки платформы смарт-контрактов.

СПИСОК ЛИТЕРАТУРЫ:

1. «Умные» контракты и современное договорное право. [Электронный ресурс]. — Режим доступа к ресурсу: https://zakon.ru/blog/2016/8/12/umnye_kontrakty_i_sovremennoe_dogovornoe_pravo. (Дата обращения: 12.12.16)
2. Smart Contracts: The Blockchain Technology That Will Replace Lawyers. [Электронный ресурс]. — Режим доступа к ресурсу: http://blockgeeks.com/guides/smart-contracts/ (Дата обращения: 09.11.2016)
3. Blockchain [Электронный ресурс]. — Режим доступа к ресурсу: http://www.investopedia.com/terms/b/blockchain.asp (Дата обращения: 27.10.2016)
4. The world is now open for business. [Электронный ресурс]. — Режим доступа к ресурсу: https://www.blockchain.com (Дата обращения: 17.01.17)
5. Перевернуть мир. Разработка уникального решения на основе технологии блокчейн. [Электронный ресурс]. — Режим доступа к ресурсу: https://yadi.sk/i/6Axqh0-ipjdHw (Дата обращения: 09.12.2016)
6. Smart Contracts Explained [Электронный ресурс]. — Режим доступа к ресурсу: http://www.blockchaintechnologies.com/blockchain-smart-contracts (Дата обращения: 13.12.2016)
7. Умные контракты (Четвертая революция стоимости). [Электронный ресурс]. — Режим доступа к ресурсу: http://old.computerra.ru/1998/266/194332/ (Дата обращения: 21.01.17)
8. Умные контракты помогут государствам принять криптовалюты. [Электронный ресурс]. — Режим доступа к ресурсу: http://ensrationis.com/умные-контракты-помогут-государства/ (Дата обращения: 23.10.16)
9. Ahmed Kosba. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts/ Ahmed Kosba, Andrew Miller, Elaine Shi, Zikai Wen, Charalampos Papamanthou; University of Maryland. — 2016. — С. 1-31
10. Guy Zyskind. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts/ Guy Zyskind, Oz Nathan, Alex Pentland; MIT Living Lab. — 2015. — С. 1-10
11. Венбо, Мао Современная криптография: теория и практика, 2005. — 677 с.
12. Enigma: Decentralized Computation Platform with Guaranteed Privacy [Электронный ресурс]: — Режим доступа к ресурсу: http://enigma.media.mit.edu/enigma_full.pdfpdf (Дата обращения: 10.10.2016)
13. Blockchain Enigma. Paradox. Opportunity [Электронный ресурс]. — Режим доступа к ресурсу: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/Innovation/deloitte-uk-blockchain-full-report.pdf (Дата обращения 17.12.16)
14. Distributed Hash Tables [Электронный ресурс]. — Режим доступа к ресурсу: https://www.cs.cmu.edu/~dga/15-744/S07/lectures/16-dht.pdf (Дата обращения: 12.01.17)
15. Децентрализованные сети: Kademlia DHT protocol [Электронный ресурс]. — Режим доступа к ресурсу: https://habrahabr.ru/post/107342/ (Дата обращения: 14.12.2016)
16. Shamir's Secret Sharing Scheme [Электронный ресурс]. — Режим доступа к ресурсу: http://point-at-infinity.org/ssss/ (Дата обращения: 20.11.2016)
17. Blockchain and Health IT [Электронный ресурс]. — Режим доступа к ресурсу: https://www.healthit.gov/sites/default/files/1-78-blockchainandhealthitalgorithmsprivacydata_whitepaper.pdf (Дата обращения: 23.01.17)

REFERENCES:

[1]. "Smart" contracts and modern contractual right [Electronic resource]. — Resource access mode: https://zakon.ru/blog/2016/8/12/umnye_kontrakty_i_sovremennoe_dogovornoe_pravo (Date of reference: 12.12.16) (in Russian).
[2]. Smart contract [Electronic resource]. — Resource access mode: http://blockgeeks.com/guides/smart-contracts/ (Date of reference: 09.11.2016)

[3]. Blockchain [Electronic resource]. — Resource access mode: http://www.investopedia.com/terms/b/blockchain.asp (Date of reference: 27.10.2016)
[4]. The world is now open for business [Electronic resource]. — Resource access mode: https://www.blockchain.com (Date of reference: 17.01.17)
[5]. To change the world. Development of unique solutions based on blockchain technology [Electronic resource]. — Resource access mode: https://yadi.sk/i/6Axqh0-ipjdHw (Date of reference: 09.12.2016) (in Russian).
[6]. Smart Contracts Explained [Electronic resource]. — Resource access mode: http://www.blockchaintechnologies.com/blockchain-smart-contracts (Date of reference: 13.12.2016)
[7]. Smart contracts (Fourth revolution value). [Electronic resource]. — Resource access mode: http://old.computerra.ru/1998/266/194332/ (Date of reference: 21.01.17) (in Russian).
[8]. Smart contracts will enable States to accept the cryptocurrency. [Electronic resource]. — Resource access mode: http://ensrationis.com/умные-контракты-помогут-государства/ (Date of reference: 23.10.16) (in Russian).
[9]. Ahmed Kosba. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts/ Ahmed Kosba, Andrew Miller, Elaine Shi, Zikai Wen, Charalampos Papamanthou; University of Maryland. — 2016. — P. 1-31
[10]. Guy Zyskind. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts/ Guy Zyskind, Oz Nathan, Alex Pentland; MIT Living Lab. — 2015. — C. 1-10
[11]. Mao V. Modern Cryptography: Theory and Practice. 2005. 677 p.
[12]. Enigma: Decentralized Computation Platform with Guaranteed Privacy [Electronic resource]: — Resource access mode: http://enigma.media.mit.edu/enigma_full.pdfpdf (Date of reference: 10.10.2016).
[13]. Blockchain Enigma. Paradox. Opportunity [Electronic resource]. — Resource access mode: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/Innovation/deloitte-uk-blockchain-full-report.pdf (Date of reference: 17.12.16).
[14]. Distributed Hash Tables [Electronic resource]. — Resource access mode: https://www.cs.cmu.edu/~dga/15-744/S07/lectures/16-dht. (Date of reference: 12.01.17).
[15]. Decentralized network. Kademlia dht protocol [Electronic resource]. — Resource access mode: https://habrahabr.ru/post/107342/ (Date of reference: 14.12.2016) (in Russian).
[16]. Shamir's Secret Sharing Scheme [Electronic resource]. — Resource access mode: http://point-at-infinity.org/ssss (Date of reference: 20.11.2016).
[17]. Blockchain and Health IT [Electronic resource]. — Resource access mode: https://www.healthit.gov/sites/default/files/1-78-blockchainandhealthitalgorithmsprivacydata_whitepaper.pdf (Date of reference: 23.01.17).

Поступила в редакцию – 26 января 2017 г. Окончательный вариант – 20 мая 2017 г.
Received – January 26, 2017. The final version – May 20, 2017.

Филипп Б. Буртыка
Южный федеральный университет, ул. Большая Садовая, 105/42, г. Ростов-на-Дону, 344006, Россия, e-mail: bbfilipp@ya.ru, ORCID iD 0000-0002-7879-802X

МЕТОДЫ ОРГАНИЗАЦИИ ПРОИЗВОЛЬНЫХ КОНФИДЕНЦИАЛЬНЫХ ЗАПРОСОВ К КРИПТОГРАФИЧЕСКОЙ ОБЛАЧНОЙ БАЗЕ ДАННЫХ С ПОМОЩЬЮ ПОЛНОСТЬЮ ГОМОМОРФНОГО ШИФРОВАНИЯ

DOI: http://dx.doi.org/10.26583/bit.2017.2.03

Аннотация. Обработка запросов к зашифрованной базе данных без ее расшифрования в последнее время активно исследуется как специалистами по криптографии, так и исследователями в области БД. Такую обработку позволяют проводить различные виды т.н. обрабатываемого шифрования (ОШ), а также специальные архитектуры систем управления базами данных (СУБД), использующие эти виды шифрования. Самыми известными видами ОШ являются шифрование, сохраняющее порядок, гомоморфное шифрование, функциональное шифрование, поисковое шифрование и шифрование, сохраняющее свойства. На основе этих видов шифрований строятся различные СУБД, наиболее известными примерами которых могут служить CryptDB, Monomi, Arx и СУБД исследователей из Новосибирска. Эти СУБД строятся на применении комбинации различных видов ОШ, например, шифрования, сохраняющего порядок, гомоморфного шифрования и традиционных видов шифрования. Однако, такой подход может вызвать проблемы с конфиденциальностью. Наилучшим подходом с точки зрения защищенности было бы построение криптографической СУБД с использованием только гомоморфного шифрования. Препятствием к этому пока что является недостаточная эффективность самого гомоморфного шифрования, а также неполное решение комплекса вопросов, связанных с необходимостью конфиденциального принятия решения в недоверенной среде. В данной работе предлагаются методы для решения этих вопросов, в частности для организации выполнения произвольного конфиденциального запроса к зашифрованной реляционной БД с помощью ПГШ. Также предлагается модель запроса к БД, разделяющая условие запроса на атомарные предикаты и связывающее условие. Один из предлагаемых методов направлен на сохранение конфиденциальности связывающего условия, другие на сохранение конфиденциальности атомарных предикатов. Параметры предложенных методов позволяют реализовать их с использованием уже существующих схем гомоморфного шифрования. Предлагаемые методы могут служить основой для построения систем защищенных криптографических облачных баз данных.

Ключевые слова: конфиденциальные запросы, криптографическая база данных, полностью гомоморфное шифрование, защищенные вычисления, секретные вычисления, недоверенная среда, криптографический протокол, конфиденциальное получение информации

Для цитирования. **БУРТЫКА, Филипп Б.** Методы организации произвольных конфиденциальных запросов к криптографической облачной базе данных с помощью полностью гомоморфного шифрования. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 24-34, June 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/102>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.03.

Filipp B. Burtyka
Southern federal university, Bolshaya Sadovaya St., 105/42 Rostov-on-Don, 344006, Russia, e-mail: bbfilipp@ya.ru, ORCID iD 0000-0002-7879-802X

The Techniques for Arbitrary Secure Quering to Encrypted Cloud Database Using Fully Homomorphic Encryption

DOI: http://dx.doi.org/10.26583/bit.2017.2.01

Abstract. The processing of queries to an encrypted database without its decryption has been actively investigated recently by both cryptographers and database researchers. Such a work is allowed by various types of so-called Processable encryption (PE), as well as special architectures of database management systems (DBMS) which use these types of encryption. The most known types of PEs are order-preserving encryption, homomorphic encryption, functional encryption, searchable encryption, and property-preserving encryption. Based on these types of encryption the various DBMSs are built, the most famous of which are CryptDB, Monomi, Arx and DBMS by researchers from Novosibirsk. These DBMSs are built on the basis of various types of PEs, for example order-preserving encryption, homomorphic encryption and traditional block encryption. However this approach can cause privacy problems. The best approach from the security viewpoint is to build a cryptographic database using only homomorphic encryption. An obstacle to this is insufficient efficiency of the existing homomorphic encryption schemes and incomplete solution of a set of issues related to ensuring the confidentiality of decisions making in an untrusted environment. In this paper we propose the techniques for solving these problems, in particular for organization of execution arbitrary secure query to the encrypted relational database using fully homomorphic encryption. Also we propose a model of query condition that splits query into atomic predicates and linking condition. One of reposed technique is aimed at ensuring the security of linking condition of queries, others keep security of atomic predicates. The parameters of the proposed techniques make it possible to implement them using the already existing homomorphic encryption schemes. The proposed techniques can be a basis for building secure cryptographic cloud databases.

Keywords: Private Queries, Encrypted Database, Fully Homomorphic Encryption, Secure Computations, Secret Computations, Untrusted Environment, Cryptographic Protocol, Private Information Retrieval, Querying Encrypted Data

For citation. **BURTYKA, Filipp B.** The Techniques for Arbitrary Secure Quering to Encrypted Cloud Database Using Fully Homomorphic Encryption. IT Security (Russia), [S.l.], v. 24, n. 2, p. 24-34, June 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/102>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.03.

На сегодняшний день облачные вычисления представляют собой ситуацию недоверенной вычислительной среды: нарушитель безопасности может быть как в канале передачи данных, так и на стороне сервера – за 2013 год организацией “Online Trust Alliance” (“Сообщество доверия в Сети”) зафиксировало около 740 млн. инцидентов безопасности [1], ещё большее число осталось незамеченным, и год от года это количество неуклонно возрастает. Единственным противодействием в этой ситуации является полное шифрование данных без расшифрования их вне клиентского компьютера. Вместе с тем естественным является желание использовать облачные ресурсы не только как хранилище данных, но и как вычислительные мощности. Таким образом, необходимо уметь проводить все необходимые вычисления над зашифрованными данными.

Инструментом для этого является т.н. обрабатываемое шифрование (ОШ, англ. processable encryption, PE) которое позволяет проводить над зашифрованными данными те или иные манипуляции без расшифрования. К такому типу шифрования относятся: гомоморфное шифрование (ГШ), шифрование сохраняющее порядок (ШСП, англ. order-preserving encryption, OPE), функциональное шифрование (ФШ, англ. functional encryption), поисковое шифрование (ПШ), шифрование, сохраняющее свойства (ШСС) и ряд других. Рассмотрим подробнее каждый из этих видов шифрования.

Суть ГШ состоит в том, что на множестве открытых текстов рассматривается одна или более операции [2], каждой из которых соответствует операция на множестве шифртекстов так, что при проведении операций над шифртекстами после расшифрования результат получается идентичным тому, как если бы соответствующие операции

проводились бы над открытыми данными. Частным случаем гомоморфного шифрования является полностью гомоморфное шифрование (ПГШ), которое гомоморфно по набору операций над открытыми данными, позволяющему выразить любую арифметическую или логическую функцию и, следовательно, теоретически провести любые вычисления над зашифрованными данными без их расшифрования [3]. Механизм действия ШСП [4] состоит в том, что на множестве шифртекстов вводится отношение порядка, «изоморфное» отношению порядка на множестве открытых текстов. ФШ характеризуется тем, что [5] позволяет вычислять некоторые заданные специальным ключом функции от зашифрованных данных, при этом результат вычислений может быть представлен в открытом виде. ПШ [6] позволяет зашифровать набор текстовых документов так, что для этого зашифрованного набора возможно проводить поиск документов, содержащих заданное зашифрованное ключевое слово.

Перейдем теперь к рассмотрению того, как эти виды шифрования используются для защиты облачных вычислений. Одними из самых распространенных типов вычислений в облаках являются алгоритмы работы с реляционными базами данных (см., например, сервис Amazon RDS [7]), такие как условные запросы (например, SQL-запросы). Выполнение такого запроса может, например, заключаться в том, что из БД выбираются записи, удовлетворяющие некоторым критериям, например, у которых заданный атрибут лежит в фиксированном диапазоне или точно равен какому-то значению. Запрос такого вида называется неагрегированным, что означает, что длина ответа на этот запрос не фиксирована (т.е. может содержать произвольное количество значений, зависящее от условий). Помимо неагрегированных запросов к БД, должна быть возможность проводить и агрегированные вычислительные запросы, такие как вычисление среднего или минимального/максимального значения по некоторому атрибуту A_i, в которых объем возвращаемых данных не зависит от содержимого БД.

Модель условия запроса

Ключевым моментом защиты облачной СУБД является обеспечение конфиденциальности запроса. В данном разделе предлагается модель декомпозиции условия запроса к БД и анализ защищенности существующих криптографических решений с использованием этой модели.

Пусть $\mathcal{R}(A_1, \dots, A_\tau))$ – реляционная схема [8] степени τ, которой соответствует база данных $\mathcal{D}$. К базе данных $\mathcal{D}$, состоящей из набора кортежей $(\alpha_1, \dots, \alpha_n)$, возможны запросы на добавление нового кортежа, либо запросы на извлечение кортежей. В последнем случае используется оператор Q, извлекающий из $\mathcal{D}$ все кортежи $(\alpha_{i_1}, \dots, \alpha_{i_{n'}})$, удовлетворяющие некоторому условию на атрибуты A_i. На SQL-подобном языке можно записать такой запрос как

SELECT * FROM $\mathcal{R}$ WHERE condition (1)

Условие condition может быть записано в виде

$\mathcal{B}(\mathcal{P}_1(A_1), \dots, \mathcal{P}_\tau(A_\tau))$, (2)

где $\mathcal{P}_i$ – *атомарные предикаты* для атрибутов A_i, а $\mathcal{B}$ – *связывающее условие*, булева формула от τ переменных. Условие вида (2) встречается и в агрегированных запросах вида

SELECT COUNT(*) FROM $\mathcal{R}$ WHERE condition (3)

Теперь рассмотрим ситуацию, когда БД находится в зашифрованном виде на недоверенном сервере, т.е. сервер хранит набор зашифрованных кортежей $(\overline{\alpha}_1, \dots, \overline{\alpha}_n)$ (здесь и далее по тексту будем обозначать как $\overline{m}$ зашифрованное значение элемента данных m, при этом предполагается, что сложные структуры данных шифруются поразрядно), и к ней необходимо *произвести запросы типа (1),(3) не производя*

расшифрования ни на каком этапе в процессе получения ответа (отметим, что запросы на добавление записей в зашифрованную БД не вызывают проблем с конфиденциальностью, вопрос же удаления записей выходит за рамки рассмотрения данной работы). Дополнительно можно ставить задачу скрытия параметров запроса Q т.е. атомарных предикатов $\mathcal{P}_i$ и/или связывающего условия $\mathcal{B}$ или, другими словами, невозможности узнавания этих параметров без ключа расшифрования.

Существует много работ, посвященных решению задач организации обработки зашифрованной реляционной БД в недоверенной среде без расшифрования с использованием различных видов ОШ.

Условно эти работы можно разделить на две группы. К первой относятся работы практической направленности, т.е. имеющие программную реализацию, работающую за приемлемое время для БД промышленных размеров (более 2^{30} записей). Сюда можно отнести CryptDB [9], MONOMI [10], программный комплекс, разработанный исследователями из Новосибирска [11]. Во всех этих работах предлагается следующее решение: разные атрибуты шифровать разными шифрами, при этом выбор шифра, применяемого к атрибуту A_i, определяется видом атомарного предиката $\mathcal{P}_i$, который необходимо будет вычислять над данными этого атрибута. Для неагрегированных запросов вида (1) в [9-11] предполагают, что предикат $\mathcal{P}_i(A_i)$ является либо предикатом сравнения на равенство вида $A_i = a$, либо предикатом сравнения порядка вида $A_i > a, A_i < a$. Для работы с предикатами сравнения на равенство используются детерминированные блочные шифры (такие как DES, AES, ГОСТ), а для предикатов сравнения порядка используется ШСП. Если же необходимо обрабатывать агрегированные арифметические запросы (сумма, среднее, дисперсия и т.д.) для атрибута A_i, то к нему применяют гомоморфное шифрование. В случае если по одному атрибуту нужно уметь проводить запросы с разными видами предикатов, то значения этого атрибута дублируются, а дубли шифруются разными шифрами.

Отметим, что в случае использования детерминированного шифрования и ШСП, неагрегированные запросы к БД выполняются тривиально т.к. клиент зашифровывает сами данные a, но не шифрует ни атомарные предикаты, ни связывающее условие $\mathcal{B}$. В этом случае, хотя сервер не знает данных, лежащих в БД, и значения a, он получает результат вычисления предикатов $\mathcal{P}_i(A_i)$ в незашифрованном виде и использует его для принятия решения о соответствии данной записи параметрам запроса. Ясно, что в этом случае при проведении запроса недоверенная сторона получит некоторую информацию, поскольку узнает, сколько записей удовлетворяют запросу клиента.

Данная общая концепция в указанных работах реализуется немного по-разному. В частности, в CryptDB [9] используется сочетание трёх видов шифрования: блочного AES, аддитивно гомоморфного шифрования Пэйе и варианта ШСП, предложенного авторами [9], которое было ноу-хау проекта, поскольку, по словам разработчиков, достигает максимума возможной защищенности для такого типа криптосистем. В работе же [11] используются только шифры, разработанные самими авторами.

В целом, такой подход, однако, ведет к проблеме с безопасностью, поскольку в этих условиях становятся возможны некоторые специфические, но довольно эффективные виды атак [12], [13]. Еще одним недостатком является то, что ШСП было признано неспособным обеспечить необходимый уровень конфиденциальности в принципе [3].

Работы второй категории имеют лишь теоретическое значение, хотя иногда и предлагают некоторые экспериментальные «proof of concept» реализации. В работах [14] и [15] были предложены способы организации работы с зашифрованными ячейками памяти с произвольным доступом (в нашем контексте то же самое что БД) при выполнении алгоритмов с зависимостями по данным (что необходимо для принятия решения по зашифрованному условию) с использованием так называемого многоаргументного

функционального шифрования (МФШ). Однако это решение обладает рядом недостатков: в первую очередь, необходимо иметь действующий экземпляр такой криптосистемы (которого на текущий момент нет), однако даже в случае достаточно высокой криптостойкости самого МФШ существуют угрозы от злонамеренных функциональностей, т.е. возможность узнать с помощью специального ключа критические сведения о данных, позволяющие раскрыть их (дешифровать). Далее, подготовительная работа, производимая клиентом перед запуском вычислений в недоверенной среде по сложности сопоставима с самим вычислением, таким образом, может дать выигрыш производительности для клиента только в очень специфических случаях (например, экспоненциальное количество запусков одной и той же функциональности над разными наборами данных).

Большую перспективу имеют решения с использованием ПГШ, однако параметры существующих полностью гомоморфных криптосистем сдерживают практическое применение этих решений. Было предпринято много попыток преодоления трудностей связанных с системами защиты БД, основанных на использовании ПГШ, и созданию и оптимизации методов и формул, реализующих эффективно реализующих необходимую для БД функциональность.

Одной из первых работ на эту тему была [16], в которой анализировалось как использовать ПГШ для организации запросов к БД, выходящих за пределы агрегатных функций и численных вычислений, таких как выборка, запросы по диапазону и соединенные. Было показано, что с использованием полностью гомоморфного шифрования, которое позволяет складывать и умножать биты открытого текста в зашифрованном виде возможно осуществлять сложные неагрегированные запросы в недоверенной среде без расшифрования, такие как выборка по диапазону значений, соединение таблиц, группировки. В этом случае на одном из этапов данные гомоморфно шифруются, а ответ возвращается клиенту в т.н. буфере результатов. При этом для неагрегированных запросов не гарантированно, что все результаты смогут быть восстановлены из буфера, хотя это может быть сделано с большой вероятностью. В числе информации, открытой для недоверенной стороны остается вычисляемый предикат, используемый в качестве условия, а также верхняя оценка на количество данных соответствующих неагрегированным запросам.

В работе [17] были предложены похожие решения. Однако протокол для неагрегированных запросов из [17] имеет некоторые особенности. На первом этапе сервер вычисляет гомоморфно булеву формулу $\mathcal{B}\left(\mathcal{P}_1(A_1),\dots,\mathcal{P}_1(A_\tau)\right)$ для каждой строки таблицы. Результатом является $\overline{1}$ или $\overline{0}$. Затем сервер гомоморфно складывает полученные значения и отправляет клиенту зашифрованное число $\overline{n}$ записей таблицы, соответствующих запросу. На втором этапе сервер получает от клиента в открытом виде верхнюю оценку n' на число n ($n' \geq n$) и возвращает (после вычислений и гомоморфной сортировки) n' записей, удовлетворяющих запросу. Такой протокол дает гарантию, что все результаты будут получены клиентом, однако имеет уязвимости, позволяющие проводить некоторые специфические атаки.

В работе [18] был рассмотрен вопрос представления алгоритмов обработки запроса над зашифрованными с помощью ПГШ данными в виде набора схем из функциональных элементов (СФЭ). В случае ПГШ над битами это будут обычные комбинационные схемы. В [18] описаны базовые варианты СФЭ для различных видов запросов и приведены их оптимизации. При этом оптимизации проводились заточенные под использование так называемого пакетированного варианта ПГШ, позволяющего проводить вычисления параллельно над несколькими наборами открытых данных. Рассматривались конъюнктивные, дизъюнктивные и пороговые запросы. Тип запроса не скрывался. В качестве предикатов рассматривались только сравнения на равенство. Для

неагрегированных запросов в [18] сервер отправлял клиенту всю зашифрованную таблицу целиком. В этой таблице значения атрибутов записей, не соответствующих запросу превращались в ноль посредством умножения всех значений записи на результат гомоморфного применения к этой записи булевой формулы (2). Т.е коммуникационная сложность была эквивалентна отсылке всей БД и оптимизация подобная [16] и [17] не производилась.

В работе [19] были более подробно проанализированы схемные сложности предложенных СФЭ, приведены результаты экспериментов, а также был предложен интересный протокол для обработки неагрегированных запросов. Этот протокол включал интерполяцию полинома, подобно протоколу из [20], что позволило избежать вычислительно трудоемкой гомоморфной сортировки. Данное решение предполагает (как и в [16]) априорное знание клиентом количества записей в БД, соответствующих запросу. С этой целью был предложен следующий механизм: для каждого типа запроса и каждого набора параметров, который интересует клиента, он заводит у себя ячейку в которой хранит число – верхнюю оценку количества записей, подходящее под этот запрос с указанными параметрами. Время от времени клиент посылает на сервер запросы на вычисление в зашифрованном виде этой оценки и сохраняет в ячейку максимальное значение. В случае необходимости произвести запрос с данным набором параметров клиент извлекает из ячейки это значение верхней оценки и отправляет его на сервер в составе запроса. Недостатком данной системы является то, что клиенту необходимо запоминать для каждого запроса значение оценки, что может потребовать больших ресурсов памяти в случае, если клиенту требуется вычислять достаточно разнообразные запросы.

В статье [21] рассматриваются дополнительные оптимизации СФЭ для агрегированных и неагрегированных запросов из [18], в частности используются два метода оптимизации – первый позволяет реализовать в пакетном режиме единичные базовые операции, а второй позволяет оптимизировать СФЭ для целочисленного сложения даже для случая меньших чем число пространств открытых текстов. Для СФЭ проверки равенства были предложены оптимизации, уменьшающие мультипликативную глубину.

В работе [22] также рассматривалась задача организации запросов по диапазону с использованием ПГШ. Была предложена реализация с использованием ПГШ из [23], и проведены эксперименты для разных значений параметров уровня криптостойкости.

Авторами [24] был предложен и проанализирован фреймворк для динамической интерпретации и компиляции запросов в булевы СФЭ. Также был рассмотрен вопрос автоматической оптимизации булевых СФЭ, получаемых в процессе преобразования программного кода и проведены эксперименты для оценки эффективности фреймворка.

В работе [25] были предложены протоколы для конфиденциального вычисления трех видов запросов конъюнктивного, дизъюнктивного и порогового, причем впервые предлагалось зашифровывать от недоверенного сервера то, с каким именно из этих трех типов запросов обратился клиент.

Однако ни в какой из имеющихся работ не была рассмотрена задача организации *произвольного* конфиденциального запроса к зашифрованной гомоморфной криптосистемой базе данных и задача полного зашифрования всех параметров запроса.

Метод обеспечения конфиденциальности произвольного запроса к зашифрованной базе данных

Как уже было сказано, впервые возможность скрытия с помощью шифрования типа запроса к полностью гомоморфно зашифрованной БД обсуждалась в работе [25], однако авторы ограничились возможностью делать один из трех типов запросов – пороговый,

конъюнктивный и дизъюнктивный (т.е. в нашей терминологии речь идет об ограничении на связывающее условие – оно может быть только конъюнкцией всех переменных, дизъюнкцией всех переменных, либо пороговым условием, принимающем истинное значение тогда и только тогда когда количество истинных переменных превышает пороговое значение). Однако, что если мы хотим делать к зашифрованной реляционной БД запросы достаточно произвольного вида и использовать шифрование для сохранение конфиденциальности вида запроса? В данном разделе предлагается метод для использования произвольного связывающего условия.

Любая булева функция от τ переменных может быть выражена в виде полинома Жегалкина (булева многочлена) [26]:

$$\sum_{i=0}^{2^{\tau}} \beta_i \cdot x_0^{i_0} \cdot \ldots \cdot x_{\tau-1}^{i_{\tau-1}}, \quad (4)$$

где $i_0, \ldots, i_{\tau-1} \in \{0,1\}$ – разряды битового представления числа i (i_0 – самый младший разряд), а β_i – коэффициенты, изменяя которые можно получить всевозможные функции. В зависимости от коэффициентов его можно превратить в любой полином от τ переменных.

Рассмотрим пример для трех переменных x_0, x_1, x_2. Для трех переменных формула (4) примет вид

$$\beta_0 \oplus \beta_1 x_0 \oplus \beta_2 x_1 \oplus \beta_3 x_2 \oplus \beta_4 x_0 x_1 \oplus \beta_5 x_1 x_2 \oplus \beta_6 x_0 x_2 \oplus \beta_7 x_0 x_1 x_2. \quad (5)$$

Пороговый запрос для трех переменных, срабатывающий при двух и более переменных выражается полиномом Жегалкина

$x_0 x_2 \oplus x_1 x_2 \oplus x_0 x_1$,

что соответствует формуле (5) при наборе коэффициентов $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_7 = 0$, $\beta_4 = \beta_5 = \beta_6 = 1$. Пороговый запрос для трех переменных, срабатывающий при одной и более переменных совпадает с дизъюнкцией трех переменных и имеет вид

$x_0 \oplus x_1 \oplus x_2 \oplus x_0 x_1 \oplus x_1 x_2 \oplus x_0 x_2 \oplus x_0 x_1 x_2$,

который может быть получен из (5) при наборе коэффициентов $\beta_0 = 0$, $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7$. Пороговый запрос, срабатывающий только при трех переменных, совпадает с конъюнкцией трех переменных и выглядит как

$x_0 x_1 x_2$,

что соответствует формуле (5) с набором коэффициентов $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ и $\beta_7 = 1$.

Основной идеей предлагаемого метода является то, что в случае, когда коэффициенты β_i зашифрованы, невозможно определить, какая именно функция вычисляется. Таким образом, подставляя вместо переменных x_i зашифрованные результаты вычисления предикатов и гомоморфно вычисляя формулу (4) с использованием зашифрованных коэффициентов $\overline{\beta}_i$ недоверенный вычислитель получает зашифрованный бит, обозначающий выполнимость условия запроса для данной зашифрованной записи.

Оценим мультипликативную глубину [27] схемы вычисления полинома (4) а также общий объем вычислений. В формуле (4) моном наибольшей степени – это произведение τ переменных предикатов и переменной-коэффициента β_i, а, как известно [27], мультипликативная глубина СФЭ, вычисляющей произведение $\tau + 1$ переменных, будет $\lceil \log_2(\tau + 1) \rceil$. Следовательно, это и будет мультипликативной глубиной схемы.

Поскольку полином (4) является функцией симметрической по переменным $x_0, \ldots, x_{\tau-1}$, то можно воспользоваться построением из [27] добавив на выход схемы лишь домножения на коэффициенты β_i и в итоге получив общую мультипликативную сложность СФЭ порядка $\tau + 3\sqrt{\tau} + 2^{\tau}$.

Коммуникационная же сложность метода составляет $O(2^{\tau})$, поскольку в ходе выполнения запроса клиент должен переслать серверу 2^{τ} зашифрованных коэффициентов $\overline{\beta}_i$. Таким образом, предлагаемый метод можно легко применять при количестве атрибутов в БД до 25-30.

Обеспечение конфиденциальности типа предиката и его параметров

Отдельно рассмотрим скрытие типа предиката и его параметров. В общем случае, поскольку речь идет о конкретной компьютерной системе, каждый из атрибутов A_i может принимать лишь конечное число значений из своей области определения $\mathrm{dom}\, A_i$. При этом предикат $\mathcal{P}_i$ задает некоторое конечное подмножество в $\mathrm{dom}\, A_i$, а размер этого подмножества изменяется от одного элемента до $|\mathrm{dom}\, A_i|$ ($|\mathrm{dom}\, A_i|$ обозначает размер множества $\mathrm{dom}\, A_i$).

Частный случай, когда предикат определялся на числовом множестве и являлся сравнением на равенство и больше/меньше рассматривался в работах [28],[29]. При этом константы, с которыми сравнивалось проверяемое в предикате значение, задавались в зашифрованном виде. Заметим, что не только числовую, но и любую другую конечную область можно представить в виде числового множества, отождествив ее с числами от 1 до $|\mathrm{dom}\, A_i|$. Например, условие, определяющее длину текстовой строки, может быть выражено как диапазон значений строк. Таким образом, если подмножество, задаваемое предикатом $\mathcal{P}_i$, представляет собой непрерывный интервал на отрезке $[1, |\mathrm{dom}\, A_i|]$, то его можно задать с помощью двух констант, при этом для определения принадлежности этому интервалу использовать алгоритм компаратора [28], который, принимая на вход два целочисленных значения одинаковой разрядности, выдает истину, если первое значение больше второго. Алгоритм компаратора для чисел заданной разрядности выражается в виде полинома Жегалкина и может быть выполнен над шифртекстами (при этом результат работы будет тоже в зашифрованном виде) [29].

В случае же если предикат должен задавать подмножество, состоящее из небольшого количества значений, лежащих далеко друг от друга, то можно воспользоваться полиномиальным представлением множества, как в [20]. При этом степень полинома определяет размер задаваемого множества элементов, которые будут корнями этого полинома. Зашифровывая его коэффициенты, тем самым делаем конфиденциальными параметры предиката.

Вышесказанное можно обобщить, составив формулу для вычисления атомарного предиката в зашифрованном виде

$$\overline{\mathcal{P}}_i = \left(\overline{\gamma}_i \oplus \overline{R}_i \oplus \overline{\gamma}_i \overline{R}_i\right)\left(\overline{\zeta}_i \oplus \overline{\delta}_i \oplus \overline{P}_i \oplus \overline{\delta}_i \overline{P}_i\right), \quad (6)$$

где $\overline{R_i}$ – зашифрованный результат вычисления диапазонного компаратора, $\overline{\gamma}_i$ – зашифрованный бит активации диапазонного компаратора, $\overline{P}_i$ – зашифрованный результат вычисления полиномиального условия, $\overline{\delta_i}$ – зашифрованный бит активации полиномиального условия, $\overline{\zeta_i}$ – зашифрованный бит инверсии полиномиального условия. Зашифрованный результат вычисления полиномиального условия равен $\overline{1}$ тогда и только тогда, когда значение представляющего множество полинома в вычисляемом значении предиката равно нулю. Зашифрованные значения $\overline{\gamma}_i, \overline{\delta}_i, \overline{\zeta}_i$ присылаются клиентом вместе с другими параметрами условия запроса.

Заключение

Представлена модель условия запроса к БД, позволяющая эффективно декомпозировать его на атомарные предикаты и связывающее условие. Предложены методы организации зашифрования атомарных предикатов и связывающего условия

запроса. Новизна методов по сравнению с известными методами обеспечения конфиденциальности запросов к криптографической БД, использующими полностью гомоморфное шифрование, заключается в том, что они дают возможность выразить любое условие на атрибуты, при этом фактически не раскрывая никакой информации об этом условии в недоверенную среду. Реализация и применение данных методов впервые делают возможным организовать произвольные конфиденциальные запросы к зашифрованной облачной базе данных эффективно в достаточной мере для практического использования и вместе с тем гарантированную защищенность исключительно криптостойкостью используемой полностью гомоморфной криптосистемы. Представленные методы могут быть реализованы с использованием любой полностью гомоморфной криптосистемы.

Более полное и математически строгое изложение будет приведено в отдельной статье.

СПИСОК ЛИТЕРАТУРЫ:

1. “Online trust alliance finds data breaches spiked to record levelin 2013”. URL: https://otalliance.org/news-events/press-releases/online-trust-alliance-finds-data-breaches-spiked-record-level-2013 (дата обращения: 25.01.2017)
2. Варновский Н. П., Шокуров А. В. Гомоморфное шифрование // Труды ИСП РАН. 2007. Вып. 12. С.27-36.
3. Бабенко Л. К., Буртыка Ф. Б., Макаревич О. Б., Трепачева А. В. Полностью гомоморфное шифрование (обзор) // Вопросы защиты информации. 2016. № 3. С. 3-25.
4. Boldyreva A., Chenette N., O’Neill A. Order-preserving encryption revisited: Improved security analysis and alternative solutions //Advances in Cryptology – CRYPTO 2011. С. 578-595.
5. Boneh D., Sahai A., Waters B. Functional encryption: Definitions and challenges // Proceedings of Theory of Cryptography Conference. – Springer Berlin Heidelberg, 2011. С. 253-273.
6. Curtmola, R., Garay, J., Kamara, S., Ostrovsky, R. Searchable symmetric encryption: improved definitions and efficient constructions //Journal of Computer Security. – 2011. – Т. 19. – №. 5. С. 895-934.
7. Amazon Relational Database Service (Amazon RDS). URL: https://aws.amazon.com/rds/ (дата обращения: 25.01.2017)
8. Дейт К. Дж. Введение в системы баз данных. М.: Вильямс, 2001. 1072 с. ISBN 5-8459-0138-3.
9. Popa, R. A., Redfield, C., Zeldovich, N., Balakrishnan, H. CryptDB: protecting confidentiality with encrypted query processing //Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles. 2011. С. 85-100. DOI: 10.1145/2043556.2043566
10. Tu, S., Kaashoek, M. F., Madden, S., Zeldovich, N. Processing analytical queries over encrypted data //Proceedings of the VLDB Endowment. – VLDB Endowment, 2013. – Т. 6. – №. 5. С. 289-300.
11. Shatilov K., Boiko V., Krendelev S., Anisutina D., Sumaneev A. Solution for secure private data storage in a cloud // Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2014, С. 885-889. DOI: 10.15439/2014F43
12. Бабенко Л.К., Трепачева А.В. Анализ защищенности одной криптографической системы для защиты конфиденциальности данных в облачных вычислениях // Безопасность информационных технологий. 2016. № 1. С. 11-15
13. Naveed M., Kamara S., Wright C. V. Inference attacks on property-preserving encrypted databases //Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, 2015. С. 644-655.
14. Goldwasser, S., Kalai, Y., Popa, R. A., Vaikuntanathan, V., Zeldovich, N. Reusable garbled circuits and succinct functional encryption //Proceedings of the forty-fifth annual ACM symposium on Theory of computing. 2013. С. 555-564. DOI: 10.1145/2488608.2488678
15. Goldwasser, S., Gordon, S. D., Goyal, V., Jain, A., Katz, J., Liu, F. H., Zhou, H. S. Multi-input functional encryption //Advances in Cryptology–EUROCRYPT 2014. 2014. С. 578-602.
16. Wang S., Agrawal D., El Abbadi A. Is Homomorphic Encryption the Holy Grail for Database Queries on Encrypted Data? // University of California, Santa Barbara, USA. 2012. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.228.4576&rep=rep1&type=pdf (дата обращения: 26.01.2017)
17. Mani M., Shah K., Gunda M. Enabling secure database as a service using fully homomorphic encryption: Challenges and opportunities //arXiv preprint 1302.2654, 2013. URL: https://arxiv.org/pdf/1302.2654 (дата обращения: 24.01.2017)

18. Cheon J. H., Kim M., Kim M. Search-and-compute on encrypted data //Proceedings of International Conference on Financial Cryptography and Data Security, 2015. С. 142-159.
19. Kim, M., Lee, H. T., Ling, S., Wang, H. On the Efficiency of FHE-based Private Queries //IEEE Transactions on Dependable and Secure Computing. – 2016.
20. Boneh, D., Gentry, C., Halevi, S., Wang, F., Wu, D. J. Private database queries using somewhat homomorphic encryption //Proceedings of International Conference on Applied Cryptography and Network Security, 2013. С. 102-118.
21. Cheon J. H., Kim M., Kim M. Optimized search-and-compute circuits and their application to query evaluation on encrypted data //IEEE Transactions on Information Forensics and Security. 2016. Т. 11, №. 1. С. 188-199. DOI: 10.1109/TIFS.2015.2483486
22. Wei S. Somewhat homomorphic encryption scheme for secure range query process in a cloud environment: диссертация на соискание степени магистра компьютерных наук. Laurentian University of Sudbury, 2015.
23. Van Dijk, M., Gentry, C., Halevi, S., Vaikuntanathan, V. Fully homomorphic encryption over the integers //Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2010. С. 24-43.
24. Stan O., Carpov S., Sirdey R. Dynamic Execution of Secure Queries over Homomorphic Encrypted Databases //Proceedings of the 4th ACM International Workshop on Security in Cloud Computing. – ACM, 2016. – С. 51-58.
25. Kim, M., Lee, H. T., Ling, S., Ren, S. Q., Tan, B. H. M., Wang, H. Better Security for Queries on Encrypted Databases. // IACR Cryptology ePrint Archive. 2016. № 470. URL: http://eprint.iacr.org/2016/470.pdf (дата обращения: 25.01.2017)
26. Марченков С. С. Замкнутые классы булевых функций. Москва: Физматлит. 2000. 128 с.
27 Boyar J., Peralta R., Pochuev D. On the multiplicative complexity of boolean functions over the basis ($\wedge, \oplus$, 1) //Theoretical Computer Science. – 2000. – Т. 235. – №. 1. – С. 43-57.
28. Togan M. A FHE-based evaluation for searching on encrypted data // International Conference on Communications (COMM), IEEE, 2016. С. 291-296.
29. Togan M., Morogan L., Plesca C. Comparison-based applications for fully homomorphic encrypted data // Proceedings of the Romanian Academy – series A: Mathematics, Physics, Technical Sciences, Information Science. 2015. Т. 16, С. 329-338.

REFERENCES:

[1]. “Online trust alliance finds data breaches spiked to record levelin 2013”. Available at: https://otalliance.org/news-events/press-releases/online-trust-alliance-finds-data-breaches-spiked-record-level-2013 (accessed 25.01.2017)
[2]. N.P. Varnovskij, A.V. Shokurov. Homomorphic encryption. Trudy ISP RAN [The Proceedings of ISP RAS], 2007. Vol. 12, pp. 27-36. (in Russian).
[3]. Babenko L. K., Burtyka F. B., Makarevich O. B. Trepacheva A.V. Fully Homomorphic Encryption (A Survey). Voprosy zashhity informacii [The information defence issues]. 2015, no. 3, pp. 3–25. (in Russian)
[4]. Boldyreva A., Chenette N., O’Neill A. Order-preserving encryption revisited: Improved security analysis and alternative solutions. Advances in Cryptology – CRYPTO 2011. pp. 578-595.
[5]. Boneh D., Sahai A., Waters B. Functional encryption: Definitions and challenges. Proceedings of Theory of Cryptography Conference. – Springer Berlin Heidelberg, 2011. pp. 253-273.
[6]. Curtmola, R., Garay, J., Kamara, S., Ostrovsky, R. Searchable symmetric encryption: improved definitions and efficient constructions.; Journal of Computer Security. 2011. Vol. 19. №. 5. pp. 895-934.
[7]. Amazon Relational Database Service (Amazon RDS). Available at: https://aws.amazon.com/rds/ (accessed on 25.01.2017)
[8]. Date C. J. An introduction to database systems. Moscow, Williams Publ., 2001, 1072 p. (in Russian)
[9]. Popa, R. A., Redfield, C., Zeldovich, N., Balakrishnan, H. CryptDB: protecting confidentiality with encrypted query processing. Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles. 2011. pp. 85-100. DOI: 10.1145/2043556.2043566
[10]. Tu, S., Kaashoek, M. F., Madden, S., Zeldovich, N. Processing analytical queries over encrypted data. Proceedings of the VLDB Endowment. – VLDB Endowment, 2013. Vol. 6. №. 5. pp. 289-300.
[11]. Shatilov K., Boiko V., Krendelev S., Anisutina D., Sumaneev A. Solution for secure private data storage in a cloud. Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2014, pp. 885-889. DOI: 10.15439/2014F43
[12]. Babenko L. K., Trepacheva A.V. Security Analysis of One Solution for Secure Private Data Storage in a Cloud. Bezopasnost' informacionnyh tehnologij [IT Security (Russia)], 2016, no. 1, pp. 11-15 (in Russian).
[13]. Naveed M., Kamara S., Wright C. V. Inference attacks on property-preserving encrypted databases. Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, 2015, pp. 644-655.

[14]. Goldwasser, S., Kalai, Y., Popa, R. A., Vaikuntanathan, V., Zeldovich, N. Reusable garbled circuits and succinct functional encryption. Proceedings of the forty-fifth annual ACM symposium on Theory of computing, 2013, pp. 555-564. DOI: 10.1145/2488608.2488678
[15]. Goldwasser, S., Gordon, S. D., Goyal, V., Jain, A., Katz, J., Liu, F. H., Zhou, H. S. Multi-input functional encryption. Advances in Cryptology – EUROCRYPT 2014, pp. 578-602.
[16]. Wang S., Agrawal D., El Abbadi A. Is Homomorphic Encryption the Holy Grail for Database Queries on Encrypted Data?; University of California, Santa Barbara, USA. 2012. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.228.4576&rep=rep1&type=pdf (accessed on: 26.01.2017).
[17]. Mani M., Shah K., Gunda M. Enabling secure database as a service using fully homomorphic encryption: Challenges and opportunities //arXiv preprint 1302.2654, 2013. Available at: https://arxiv.org/pdf/1302.2654 (accessed on: 24.01.2017).
[18]. Cheon J. H., Kim M., Kim M. Search-and-compute on encrypted data. Proceedings of International Conference on Financial Cryptography and Data Security, 2015. pp. 142-159.
[19]. Kim, M., Lee, H. T., Ling, S., Wang, H. On the Efficiency of FHE-based Private Queries. IEEE Transactions on Dependable and Secure Computing, 2016.
[20]. Boneh, D., Gentry, C., Halevi, S., Wang, F., Wu, D. J. Private database queries using somewhat homomorphic encryption. Proceedings of International Conference on Applied Cryptography and Network Security, 2013. pp. 102-118.
[21]. Cheon J. H., Kim M., Kim M. Optimized search-and-compute circuits and their application to query evaluation on encrypted data. IEEE Transactions on Information Forensics and Security, 2016, vol. 11, no. 1, pp. 188-199. DOI: 10.1109/TIFS.2015.2483486
[22]. Wei S. Somewhat homomorphic encryption scheme for secure range query process in a cloud environment: Master's thesis in Computer Science. Laurentian University of Sudbury, 2015.
[23]. Van Dijk, M., Gentry, C., Halevi, S., Vaikuntanathan, V. Fully homomorphic encryption over the integers. Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2010. pp. 24-43.
[24]. Stan O., Carpov S., Sirdey R. Dynamic Execution of Secure Queries over Homomorphic Encrypted Database. Proceedings of the 4th ACM International Workshop on Security in Cloud Computing. – ACM, 2016. – pp. 51-58.
[25]. Kim, M., Lee, H. T., Ling, S., Ren, S. Q., Tan, B. H. M., Wang, H. Better Security for Queries on Encrypted Databases. IACR Cryptology ePrint Archive, 2016, no. 470. Available at: http://eprint.iacr.org/2016/470.pdf (accessed on 25.01.2017)
[26]. Marchenkov S. S, Zamknutye klassy bulevyh funkcij. Moskva: Fizmatlit. 2000. Pp. 128.
[27]. Boyar J., Peralta R., Pochuev D. On the multiplicative complexity of boolean functions over the basis ($\wedge, \oplus$, 1). Theoretical Computer Science. – 2000. Vol. 235. №. 1. pp. 43-57.
[28]. Togan M. A FHE-based evaluation for searching on encrypted data. International Conference on Communications (COMM), IEEE, 2016, pp. 291-296.
[29]. Togan M., Morogan L., Plesca C. Comparison-based applications for fully homomorphic encrypted data. Proceedings of the Romanian Academy – series A: Mathematics, Physics, Technical Sciences, Information Science. 2015, vol. 16, pp. 329-338.

Поступила в редакцию – 21 февраля 2017 г. Окончательный вариант – 20 мая 2017 г.
Received – February 21, 2017. The final version – May 20, 2017.

Виктор С. Горбатов, Алексей А. Мещеряков
Национальный исследовательский ядерный университет «МИФИ»,
Каширское ш., 31, г. Москва, 115409, Россия,
e-mail: VSGorbatov@mephi.ru, ORCID iD 0000-0001-9998-9733
e-mail: aameshcheryakov@mephi.ru, ORCID iD 0000-0001-5148-7734

КУРС ТРЕНИНГА ПО БЕЗОПАСНОЙ РАЗРАБОТКЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

DOI: http://dx.doi.org/10.26583/bit.2017.2.04

Аннотация. Безопасность является одним из важнейших критериев качества разрабатываемого программного обеспечения. Для получения достаточного уровня защищенности разрабатываемых приложений компании внедряют процессы обеспечения безопасности в жизненных цикл разработки программного обеспечения. Так компании-разработчики сталкиваются с дефицитом кадров, способных решать задачи как по проектированию и реализации ПО, так и по обеспечению их защиты. Данная статья предлагает описание курса тренинга по безопасной разработке программного обеспечения. Курс по безопасной разработке приложений предназначен для совместного обучения студентов, получающих образование как в области проектирования и реализации ПО, так и в области информационной безопасности.

Ключевые слова: программное обеспечение, безопасность приложений, обучение, образование, информационная безопасность.

Для цитирования. **ГОРБАТОВ Виктор. С.; МЕЩЕРЯКОВ, Алексей А.** Курс тренинга по безопасной разработке программного обеспечения. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 35-41, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/103>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.04.

Viktor S. Gorbatov, Alexey A. Meshcheriakov
National Research Nuclear University MEPhI (Moscow Engineering Physics Institute),
Kashirskoe shosse, 31, Moscow, 115409, Russia,
e-mail: VSGorbatov@mephi.ru, ResearcherID B-5106-2016
e-mail: aameshcheryakov@mephi.ru, ORCID iD 0000-0001-5148-7734

Secure software development training course

DOI: http://dx.doi.org/10.26583/bit.2017.2.04

Abstract. Information security is one of the most important criteria for the quality of developed software. To obtain a sufficient level of application security companies implement security process into software development life cycle. At this stage software companies encounter with deficit employees who able to solve problems of software design, implementation and application security. This article provides a description of the secure software development training course. Training course of application security is designed for co-education students of different IT-specializations.

Keywords: information security, education, secure software development, application security, training.

For citation. **GORBATOV, Viktor. S.; MESHCHERIAKOV, Alexey A.** Secure software development training course. IT Security (Russia), [S.l.], v. 24, n. 2, p. 35-41, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/103>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.04.

Введение

Безопасность остается одним из главных приоритетов при разработке программного обеспечения (ПО) на протяжении последних лет. Однако, как показывает практика

большинство разрабатываемых приложений уязвимо к тем или иным атакам. Так исследование безопасности финансовых приложений [1], проведенное компанией Positive Technologies, показало, что все анализируемые системы дистанционного банковского обслуживания содержали уязвимости по крайней мере среднего уровня риска, при этом почти в каждой из систем (90%) были обнаружены критически опасные уязвимости. Схожую статистику по наличию уязвимостей имеют и приложения, написанные для других отраслей экономики и сфер деятельности.

Для получения приемлемого уровня защищенности разрабатываемых приложений компании внедряют процессы обеспечения безопасности в жизненных цикл разработки ПО (SDLC). Внедрение процессов безопасности приложений является сложным и длительным процессом, требующим применения специфических практик безопасности на всех этапах жизненного цикла ПО.

Проблема безопасности приложений находится на стыке ответственности программистов и специалистов по безопасности. Таким образом, от команды по разработке ПО требуется как глубокое понимание угроз и уязвимостей приложений, так и навыки реализации механизмов безопасности, анализа и тестирования кода.

Сегодня система высшего образования не готовит таких специалистов, которые бы хорошо разбирались и в области разработки ПО, и в области его безопасности. Это создает дефицит профессиональных кадров в данной области. Зачастую, компаниям-разработчикам ПО приходится самим дополнительно обучать сотрудников-программистов вопросам безопасной разработки. Сложившаяся ситуация приводит к дополнительным издержкам на разработку, а также создает сложности для донесения до всех программистов информации об актуальных угрозах и важности обеспечения безопасности.

В данной работе описывается процесс создания курса по безопасной разработке приложений для совместного обучения студентов, получающих образование как в области проектирования и реализации ПО, так и в области информационной безопасности. Совместное обучение студентов различных направлений позволяет добиться синергетического эффекта, в результате которого обучающиеся получат новые знания о смежной профессиональной области и навыки взаимодействия между командами разработки и безопасности.

Обзор литературы

Академические работы написанные в области обучения безопасности приложений, направлены преимущественно на построение процесса повышения осведомленности сотрудников в вопросах информационной безопасности и применимы для компаний-разработчиков ПО. Но многие из таких практик могут быть адаптированы для применения в учебном процессе ВУЗа.

Сообщество OWASP (Open Web Application Security Project) подготовило обучающую программу для web-разработчиков [2]. Это четырехчасовой курс, раскрывающий следующие темы: актуальные проблемы безопасности, уязвимости веб-приложений, встраивание в SDLC, обзор хороших практик безопасной разработки, тестирование безопасности.

Правительства многих стран пытается решить проблему обучения информационной безопасности. National Initiative for Cybersecurity Education (NICE) – это государственная образовательная программа кибер-безопасности США [3]. Главной целью NICE является

построение постоянное повышение уровня осведомленности в вопросах безопасности на всех уровнях образования.

В работе [4] показано, что недостаточное изучение вопросов информационной безопасности в учебных программах IT-специальностей может стать основной проблемой, приводящей к возникновению уязвимостей, и предложен подход, предполагающий проведение тренингов для повышения уровня осведомленности.

Стандарты разработки безопасных приложений

The National Institute of Standards and Technology (NIST) оценивает возрастание стоимости исправления дефекта на стадии эксплуатации ПО в 470-880 раз по сравнению с исправлением на стадии сбора требований [5]. Поэтому целесообразнее внедрять процессы обеспечения безопасности во все этапы жизненного цикла, чтобы увеличить вероятность раннего выявления дефектов безопасности.

В области разработки ПО уже подготовлен ряд корпоративных, отраслевых и международных стандартов, описывающих управление процессом создания безопасного ПО:

- ГОСТ Р ИСО/МЭК 27034 [6];
- ISO/IEC TR 24772 [7];
- Microsoft Security Development Life Cycle [8];
- Cisco Security Development Life Cycle [9];
- OpenSAMM [10];
- OWASP CLASP [11];
- BSIMM [12].

Сравнительный анализ представленных выше стандартов проведен в работе [13]. Ввиду многообразия стандартов разного уровня индустрия разработки ПО еще не выработала единого подхода, оформленного в виде единого документа. Поэтому при разработке данного курса мы выделили ряд наиболее применяемых мер и механизмов, рассмотрение которых позволит создать у слушателей целостную картину в изучаемой области:

- Обучение сотрудников безопасной разработке ПО;
- Моделирование угроз безопасности информации, источником которых является ПО;
- Определение требований по разработке безопасного ПО;
- Использование стандарта на оформление исходного кода программы (правил и рекомендаций, направленных на минимизацию количества потенциально уязвимых конструкций в исходном коде);
- Проведение статического анализа исходного кода;
- Проведение динамического анализа ПО;
- Проведение ручной экспертизы исходного кода;
- Проведение анализа уязвимостей;
- Обеспечение безопасной поставки ПО пользователю;
- Устранение уязвимостей ПО, выявленных в процессе функционирования ПО.

Учебная программа

Учебная программа курса тренинга по безопасной разработке приложений включает в себя три модуля:

1. Введение в разработку безопасного программного обучения
   a. Обзор основных угроз ПО.
   b. Безопасность web-приложений.
   c. Безопасность мобильных приложений.
   d. Безопасность десктопных приложений.
   e. Текущие тренды в области безопасности приложений.
2. Разработка безопасного ПО
   a. Формирование требований к безопасному ПО.
   b. Управление конфигурации при разработке безопасного ПО.
   c. Ручная экспертиза кода.
   d. Статический и динамический анализ кода. Фаззинг-тестирование.
   e. Тестирование на проникновение при разработке ПО.
3. Лучшие практики и приемы разработки безопасного ПО
   a. Обзор общих практик и приемов разработки безопасного ПО.
   b. Практики разработки безопасных web-приложений.
   c. Практики разработки безопасных мобильных приложений.
   d. Практики разработки безопасных десктопных приложений.

Задачей первого модуля является знакомство слушателей с актуальным состоянием проблем в рассматриваемой области. Слушатели изучают угрозы и характерные уязвимости безопасности для различных типов приложений. По итогам прохождения данного модуля у слушателя должно сложиться комплексное представление об актуальных угрозах и атаках, которым подвержены современное ПО.

Второй модуль рассматривает меры и механизмы безопасной разработки ПО. В данном модуле слушатели знакомится с моделью безопасности жизненного цикла ПО и получают теоретическую информацию об использовании различных инструментов для обеспечения безопасности приложения.

В третьем модуле изучаются лучшие практики безопасной разработки различных типов приложений. Данный модуль обобщает знания, полученные в первых разделах и дополняет их специфической информацией для каждой конкретной области.

Таким образом, учебная программа курса позволяет сформировать комплексное теоретическое представление об актуальных угрозах, уязвимостях, атаках на ПО, а также о мерах, механизмах и практиках разработки безопасных приложений.

Описание заданий

В ходе прохождения курса по разработке безопасных приложений слушатели выполняют различные домашние задания, связанные с анализом проблем защиты информации, возникающих на каждом из этапов жизненного цикла ПО. Всего курс предусматривает три задания, соответствующие теоретическим модулям:

1. Анализ и формирование требований безопасности к ПО.
2. Выявление уязвимостей в ПО.
3. Реализация модулей безопасности для ПО.

Каждое из заданий ставит своей задачей дать слушателю возможность приобрести практические навыки в области анализа безопасности приложений.

Рис. 1. Структура курса разработки безопасного ПО
Fig. 1. Course structure the secure development

В качестве курсового проекта студентам предлагается разработка какого-либо ПО, например, безопасного текстового чата. Для выполнения такого проекта по мере прохождения курса все студенты делятся на группы, в которые будут входить студенты всех IT-специальностей. Такие группы будут обладать необходимым набором компетенций для выполнения проекта по безопасной разработке. Целью разработки такого проекта является получение практических знаний и навыков в области безопасной разработки. Команды в ходе работы над проектом проходят все этапы жизненного цикла разработки ПО, используя на каждом из них необходимые инструменты обеспечения безопасности приложений. Таким образом, студенты могут оценить на практике эффективность тех или иных инструментов и методов для конкретного проекта. В конце курса каждая команда презентует свой проект перед всеми слушателями. После этого студентов предлагается проведения тестирования на проникновение приложений, разработанных другими командами.

Описание лаборатории

Промышленный процесс разработки ПО предполагает использование различных инструментов: системы управления проектом, системы контроля версий, инструменты тестирования и т.д. Для создания лаборатории были использованы такие инструменты как BurpSuite, OWASP ZAP, Redmine, Fortify [14-17]. Структурная схема развернутой лаборатории курса представлена на рисунке 2.

Рис. 2. Структурная схема лаборатории курса безопасной разработки ПО
Fig. 2. Structural scheme of the laboratory course secure software development

На базе рассмотренного выше стенда развернуты тестовые приложения с преднамеренно внесенными дефектами безопасности. На текущий момент разрабатываются следующие тестовые приложения:

- веб-приложение;
- мобильный клиент для Android.

Разрабатываемые приложения логически связаны между собой в систему имитирующую дистанционную банковскую систему (ДБО). Внесенные дефекты безопасности являются типичными для данных классов ПО и основаны на уязвимостях, найденных в реальных системах. Разрабатываемые приложения анализируются слушателями в ходе выполнения домашних заданий с целью выявления и исправления дефектов безопасности.

Заключение

В данной работе мы рассмотрели курс тренинга по безопасной разработке приложений. В данном курсе было соединено рассмотрение вопросов проектирования, реализации ПО и вопросов его безопасности. Лекции и практические задания раскрывают проблемы безопасности, характерные для различных типов приложений, и рассматривают лучшие практики безопасного программирования. Курсовой проект призван дать слушателям возможность на практике погрузиться в процесс разработки ПО, и оценить необходимость внедрения различных мер и механизмов безопасности на всех этапах жизненного цикла ПО.

Внедрение разработанного курса в учебную программу студентов высших образовательных учреждений всех IT специальностей позволит подготавливать высококвалифицированных специалистов, способных внедрять практики безопасного программирования в организациях.

СПИСОК ЛИТЕРАТУРЫ:
1. Positive Security. Уязвимости приложений финансовой отрасли. 2016. URL: https://www.ptsecurity.com/upload/ptru/analytics/Financial-Vulnerability-2016-rus.pdf (дата обращения: 22.02.2017)
2. OWASP. Education Track: What Developers Should Know on Web Application Security. URL: https://www.owasp.org/index.php/Education_Track:_What_Developers_Should_Know_on_Web_Application_Security (дата обращения: 23.02.2017)
3. Jøsang A. , Odegaard M, Oftedal E.: Cybersecurity Through Secure Software Development. 9th IFIP WG 11.8 World Conference, 2015
4. The National Initiative for Cybersecurity Education (NICE). URL: http://csrc.nist.gov/nice/ (дата обращения: 20.02.2017)
5. NIST. Creating a Patch and Vulnerability Management Program. URL: http://csrc.nist.gov/publications/nistpubs/800-40-Ver2/SP800-40v2.pdf
6. ГОСТ Р ИСО/МЭК 27034-1-2014. Информационная технология. Методы и средства обеспечения безопасности. Безопасность приложений. Часть 1. Обзор и общие понятия. М.: Стандартинформ, 2015. 64 с.
7. ISO/IEC TR 24772:2010. Information technology. Programming languages. Guidance to avoiding vulnerabilities in programming languages through language selection and use.
8. Microsoft Security Development Lifecycle. URL: https://www.microsoft.com/en-us/sdl/ (дата обращения: 15.02.2017)
9. Cisco Secure Development Lifecycle. URL: https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/cisco-secure-development-lifecycle.pdf (дата обращения: 15.02.2017)
10. Software Assurance Maturity Model. URL: http://www.opensamm.org/downloads/SAMM-1.0.pdf (дата обращения: 15.02.2017)
11. CLASP Security Principles. URL: https://www.owasp.org/index.php/CLASP_Security_Principles (дата обращения: 15.02.2017)
12. BSIMM Framwork. URL: https://www.bsimm.com/framework/ (дата обращения: 18.02.2017)
13. Barabanov A., Markov A., Fadin A., Tsirlov V., Shakhanov I. Synthesis of secure software development controls. In: SIN '15 Proceedings of the 8th International Conference on Security of Information and Networks, 2015
14. Burp Suite. URL: https://portswigger.net/burp/ (дата обращения: 21.02.2017)
15. OWASP Zed Attack Proxy Project. URL: https://www.owasp.org/index.php/OWASP_Zed_Attack_Proxy_Project (дата обращения: 21.02.2017)
16. Redmine. URL: https://www.redmine.org/ (дата обращения: 21.02.2017)
17. Fortify Static Code Analyzer. URL: http://www8.hp.com/ru/ru/software-solutions/static-code-analysis-sast/ (дата обращения: 21.02.2017)

REFERENCES:
[1] Positive Technologies. Vulnerabilities in Online Banking Systems - 2016. (In Russian. Available at: https://www.ptsecurity.com/upload/ptru/analytics/Financial-Vulnerability-2016-rus.pdf). (accessed 22.02.2017)
[2] OWASP. Education Track: What Developers Should Know on Web Application Security. (In Russian. Available at: https://www.owasp.org/index.php/Education_Track:_What_Developers_Should_Know_on_Web_Application_Security). (accessed 23.02.2017)
[3] Jøsang A. , Odegaard M, Oftedal E.: Cybersecurity Through Secure Software Development. In: 9th IFIP WG 11.8 World Conference, 2015
[4] The National Initiative for Cybersecurity Education (NICE). URL: http://csrc.nist.gov/nice/ (accessed: 20.02.2017)
[5] NIST. Creating a Patch and Vulnerability Management Program. URL: http://csrc.nist.gov/publications/nistpubs/800-40-Ver2/SP800-40v2.pdf
[6] GOST 27034-1-2014. Information technology. Security techniques. Application security. Part 1: Overview and concepts Moscow, Standartinform Publ., 2015. 64 p. (In Russian).
[7] ISO/IEC TR 24772:2010. Information technology. Programming languages. Guidance to avoiding vulnerabilities in programming languages through language selection and use.
[8] Microsoft Security Development Lifecycle. URL: https://www.microsoft.com/en-us/sdl/ (accessed: 15.02.2017)
[9] Cisco Secure Development Lifecycle. URL: https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/cisco-secure-development-lifecycle.pdf (accessed: 15.02.2017)
[10] Software Assurance Maturity Model. URL: http://www.opensamm.org/downloads/SAMM-1.0.pdf (accessed: 15.02.2017)
[11] CLASP Security Principles. URL: https://www.owasp.org/index.php/CLASP_Security_Principles (accessed: 15.02.2017)
[12] BSIMM Framwork. URL: https://www.bsimm.com/framework/ (accessed: 18.02.2017)
[13] Barabanov A., Markov A., Fadin A., Tsirlov V., Shakhanov I. Synthesis of secure software development controls. In: SIN '15 Proceedings of the 8th International Conference on Security of Information and Networks, 2015
[14] Burp Suite. URL: https://portswigger.net/burp/ (accessed: 21.02.2017)
[15] OWASP Zed Attack Proxy Project. URL: https://www.owasp.org/index.php/OWASP_Zed_Attack_Proxy_Project (accessed: 21.02.2017)
[16] Redmine. URL: https://www.redmine.org/ (accessed: 21.02.2017)
[17] Fortify Static Code Analyzer. URL: http://www8.hp.com/ru/ru/software-solutions/static-code-analysis-sast/ (accessed: 21.02.2017)

Поступила в редакцию – 20 февраля июля 2017 г. Окончательный вариант – 21 мая 2017 г.
Received – February 20, 2017. The final version – May 21, 2017.

Александр М. Коротин
Национальный исследовательский ядерный университет «МИФИ»,
Каширское шоссе, 31, г. Москва, 115409, Россия, e-mail:
e-mail: alexkor91@gmail.com, ORCID iD 0000-0003-3773-342X

АНАЛИЗ УГРОЗ БЕЗОПАСНОСТИ ОТВЕТСТВЕННОЙ ИНФОРМАЦИИ, ПЕРЕДАВАЕМОЙ СИСТЕМОЙ АЛС НА БАЗЕ РАДИОКАНАЛА

DOI: http://dx.doi.org/10.26583/bit.2017.2.05

Аннотация. Переход к использованию радиоканала для передачи ответственной информации в системах автоматической локомотивной сигнализации (АЛС) усложняет обеспечение безопасности перевозочного процесса, так как в этом случае создаются условия для получения несанкционированного доступа к каналу связи из-за пределов контролируемой зоны. В связи с этим представляется целесообразным решение задачи по определению угроз безопасности ответственной информации, передаваемой системой автоматической локомотивной сигнализации на базе радиоканала (АЛСР), реализация которых способна нарушить безопасность перевозочного процесса. Решение поставленной задачи возможно путем анализа угроз безопасности ответственной информации, определения степени их влияния на перевозочный процесс, разработки модели нарушителя. В статье показано, что нарушение безопасности перевозочного процесса возможно в случае нарушения целостности ответственной информации. Используя методические документы ФСТЭК России, разработана модель нарушителя безопасности ответственной информации. Сделан вывод, что угрозы со стороны внутреннего нарушителя для традиционных систем автоматической локомотивной сигнализации и систем АЛСР совпадают. Сформирован перечень угроз безопасности ответственной информации, направленных на нарушение безопасности перевозочного процесса со стороны внешнего нарушителя. Полученные в статье результаты могут использоваться при исследовании безопасности систем АЛСР, анализе существующих решений по защите ответственной информации, передаваемой АЛСР, или разработке новых способов защиты.

Ключевые слова: автоматическая локомотивная сигнализация на базе радиоканала, АЛСР, безопасность АЛСР, модель угроз и нарушителя АЛСР, безопасность ответственной информации, обеспечение целостности ответственной информации.

Для цитирования. **КОРОТИН, Александр М.** Анализ угроз безопасности ответственной информации, передаваемой системой АЛС на базе радиоканала. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 42-49, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/104>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.05.

Alexander M. Korotin
National Research Nuclear University MEPhI,
Kashirskoe shosse, 31, Moscow, 115409, Russia,
e-mail: alexkor91@gmail.com, ORCID iD 0000-0003-3773-342X

Security threat analysis of critical information transmitted by the ALS system based on the radio channel

DOI: http://dx.doi.org/10.26583/bit.2017.2.05

Abstract. The transition to the use of a radio channel for the transfer of critical information in automatic locomotive signaling (ALS) systems complicates the security of the transportation process, because in this case conditions are created for obtaining unauthorized access to the communication channel from outside the controlled area. That's why it seems advisable to solve the problem of identifying security threats of critical information transmitted by the automatic locomotive signaling system based on the radio channel (ALSR), the implementation of which is

likely to disrupt the safety of the transportation process. The solution of this task is possible by analyzing security threats of critical information, determining the degree of their influence on the transportation process and developing violator model. The article shows that a security breach of the transportation process is possible in case of violation of the responsible information integrity. A model of violator of critical information security was developed using methodical documents of FSTEC of Russia. It was concluded that the threats from the internal violator for traditional automatic locomotive signaling systems and ALSR systems coincide. A list of security threats of sensitive information was developed, directed to violating the security of the transportation process by an external violator. The results obtained in the article can be used to study the safety of ALSR systems, to analyze existing solutions for protecting critical information transmitted by ALSR, or to develop new methods of protection.

Keywords: automatic locomotive signaling based on radio channel, ALSR, security of ALSR, threats model and violator model of ALSR, security of critical information, ensuring of critical information integrity.

For citation. **KOROTIN, Alexander M.** Security Threat Analysis of Critical Information Transmitted by the ALS System Based on the Radio Channel. IT Security (Russia). [S.l.], v. 24, n. 2, p. 42-49, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/104>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.05.

Введение

Система автоматической локомотивной сигнализации (АЛС) представляет собой автоматизированную систему управления, применяемую на железнодорожном транспорте для обеспечения безопасности перевозочного процесса. Под безопасностью перевозочного процесса понимается такое его состояние, при котором отсутствует недопустимый риск возникновения транспортных происшествий и их последствий, причиняющих вред жизни или здоровью граждан, окружающей среде, имуществу физических или юридических лиц [1]. Одной из задач, решаемой системой АЛС в рамках обеспечения безопасности перевозочного процесса, является обмен ответственной информацией между станционными системами определения занятости пути и бортовыми устройствами безопасности [2-4]. Под ответственной информацией в статье понимается информация о допустимой скорости движения и дополнительных условиях следования железнодорожного подвижного состава (ограничения скорости, маршрут движения по железнодорожной станции), необходимая для обеспечения безопасности перевозочного процесса.

В состав системы АЛС входит станционное оборудование, расположенное на станциях или вдоль железнодорожных путей (станционная часть АЛС), и бортовое, расположенное на борту локомотива (бортовая часть АЛС) (см. рисунок 1). При этом в АЛС выделяют два логических уровня – уровень управления перевозочным процессом (прикладной уровень) и уровень передачи данных (канальный уровень). На прикладном уровне работают локомотивное устройство безопасности и шлюз взаимодействия с системой определения свободности пути [2]. Между ними происходит обмен ответственной информацией, которая используется для управления и обеспечения безопасности перевозочного процесса. На канальном уровне функционирует система связи, состоящая из напольного и бортового приемопередающего оборудования, отвечающая за доставку ответственной информации.

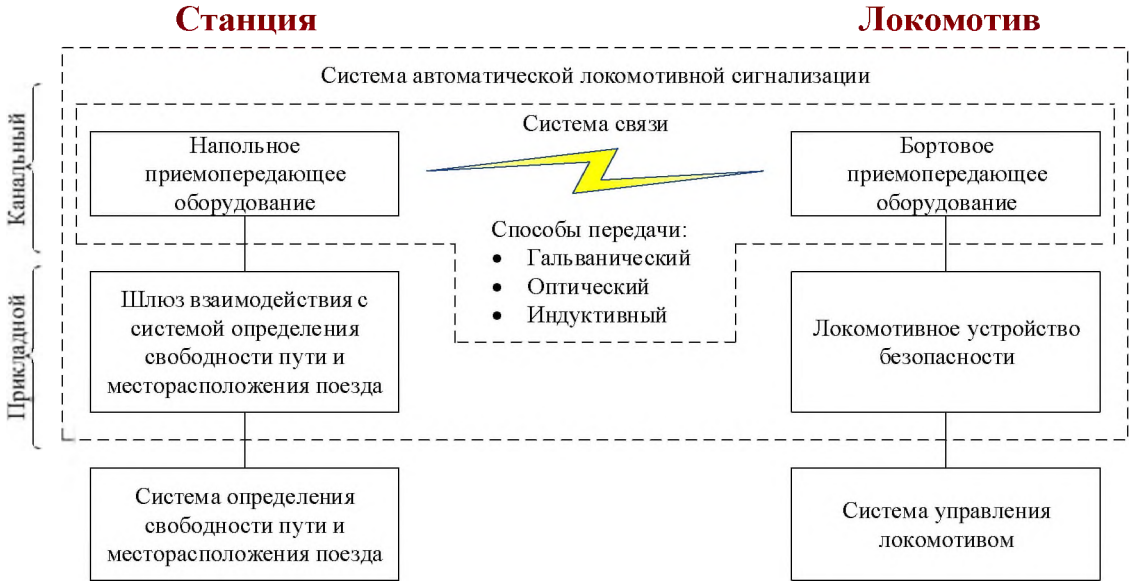


Рис. 1. Структурная схема системы АЛС
Fig. 1. Block diagram of the system ALS

Использование систем радиосвязи для передачи ответственной информации между станционной и бортовой частями АЛС обладает рядом преимуществ по сравнению с другими системами связи, например, позволяет увеличить объем передаваемых данных, уменьшить вероятность их искажения [5]. Однако, такой переход к использованию радиоканала усложняет обеспечение безопасности перевозочного процесса, так как в этом случае канал связи между станционной и бортовой частями АЛС выходит за пределы контролируемой зоны.

В связи с этим актуальной задачей является определение угроз безопасности перевозочного процесса, реализация которых возможна из-за пределов контролируемой зоны путем нарушения функции обмена ответственной информацией в системах АЛС на базе радиоканала (АЛСР). Для решения данной задачи представляется целесообразным провести анализ существующих угроз безопасности ответственной информации, передаваемой системой АЛСР. Под угрозой безопасности ответственной информации, передаваемой системой АЛСР, понимается совокупность условий и факторов, создающих потенциальную и реально существующую опасность нарушения конфиденциальности, целостности и/или доступности ответственной информации, вызывающую или способную вызвать негативные последствия (ущерб/вред) для перевозочного процесса.

Анализ влияния аспектов безопасности на перевозочный процесс

Согласно работам [6-11] по исследованию кибербезопасности железнодорожного транспорта выделяют три основных класса угроз безопасности микропроцессорных систем железнодорожной автоматики и телемеханики по степени ущерба:

- нарушение безопасности перевозочного процесса (1-ый класс);
- снижение эффективности перевозочного процесса (2-ой класс);
- угрозы, не влияющие на перевозочный процесс, однако приводящие к операционным и/или финансовым потерям (3-ий класс).

Используя данную классификацию, был выполнен анализ угроз, связанных с нарушением разных аспектов безопасности ответственной информации, передаваемой по

радиоканалу. Выявлено, что угрозы, связанные с нарушением целостности ответственной информации, способны привести к получению участниками движения неверных данных, что, в свою очередь, может стать причиной создания ложных свободных блок-участков на пути следования локомотива или изменения разрешенной скорости движения на участке. Возникновение подобных событий нарушает безопасность перевозочного процесса, следовательно, угрозы, связанные с нарушением целостности ответственной информации, относятся к первому классу угроз по уровню ущерба.

Угрозы, связанные с нарушением доступности передаваемой информации или используемого канала связи, относятся ко второму классу угроз. Создание помех в канале связи или подавление радиостанций приведет к невозможности получения ответственной информации, таким образом, подвижной состав продолжит движение по полученным ранее данным до тех пор, пока локомотивное устройство безопасности не остановит его согласно кривой торможения, заранее высчитанной локомотивным устройством безопасности. Таким образом, безопасность перевозочного процесса не будет нарушена, однако его эффективность будет снижена.

Угрозы нарушения конфиденциальности передаваемой ответственной информации могут быть отнесены к третьему классу, так как перехват и раскрытие передаваемых данных в системе АЛСР не влияют напрямую на перевозочный процесс и скорее могут быть использованы для анализа протокола передачи, определения загруженности блок-участка или частоты движения поездов в рамках подготовки проведения компьютерных атак на железнодорожную инфраструктуру.

Таким образом, был сделан вывод, что нарушение безопасности перевозочного процесса возможно только в случае реализации угроз, связанных с нарушением целостности ответственной информации.

Модель нарушителя

При анализе возможных источников угроз технические отказы аппаратуры, а также случайные (непреднамеренные) ошибки людей (персонала) не рассматривались, так как задача защиты перевозочного процесса от подобных угроз относится к области функциональной безопасности и должна выполняться самой системой АЛСР [2]. В связи с этим в качестве источников угроз в работе рассматривались только антропогенные источники – внешний нарушитель, не имеющий доступ к оборудованию системы АЛСР, и внутренний нарушитель.

При этом в качестве внутреннего нарушителя в соответствии с нормативными документами ФСТЭК России [12] рассматривались лица первой категории, имеющие доступ к оборудованию АЛСР, но не являющиеся пользователями системы (обслуживающий персонал, например, электромеханик СЦБ или ремонтная бригада [13]), второй категории, являющиеся пользователями системы (локомотивная бригада [14]), седьмой категории, являющиеся разработчиками ПО прикладного уровня и обеспечивающие его техническое сопровождение, а также восьмой категории – занимающиеся поставкой, сопровождением и ремонтом технических средств, применяемых в АЛСР. Лица 3-6 категорий в работе не рассматривались, так как процесс функционирования системы АЛСР является автономным (отсутствует управляющее воздействие со стороны персонала), и сама система не имеет привилегированных или удаленных пользователей. Кроме того, в ходе анализа было сделано предположение, что нарушители могут вступать в сговор с целью обмена имеющейся у них информацией о системе, а также консолидации возможностей по нарушению безопасности ответственной информации.

На основе выбранных источников угроз была разработана модель нарушителя безопасности ответственной информации.

Внешний нарушитель может осуществлять:

- подмену базовой и/или абонентской радиостанции;
- отправку поддельной ответственной информации в радиоканал;
- повторную отправку ранее перехваченной в радиоканале ответственной информации.

Внешний нарушитель может располагать следующей информацией о системе:

- доступные в открытых источниках сведения о системе АЛСР и ее компонентах;
- фрагменты информации, передаваемой по радиоканалу.

Внутренний нарушитель первой категории может осуществлять:

- реализацию автономных программно-аппаратных закладок в систему АЛСР на уровне ее компонентов (без внесения изменений в используемые технические средства).

Внутренний нарушитель первой категории может располагать следующей информацией:

- любая информация, доступная внешнему нарушителю;
- информация об используемых интерфейсах и способах подключения компонентов АЛСР между собой и с внешними системами;
- фрагменты информации, передаваемой между компонентами АЛСР, а также между системой АЛСР и внешними системами (системой определения свободности пути и месторасположения поезда, системой управления локомотивом и др.).

Внутренний нарушитель второй категории может осуществлять:

- все действия, доступные внутреннему нарушителю первой категории;
- внедрение вредоносных компьютерных программ в компоненты АЛСР, к которым имеется легитимный доступ.

Внутренний нарушитель первой и второй категории может располагать следующей информацией:

- любая информация, доступная внутреннему нарушителю первой категории;
- эксплуатационная документация на систему АЛСР;
- фрагменты информации, обрабатываемые компонентами системы АЛСР, к которым имеется легитимный доступ.

Внутренний нарушитель седьмой категории может осуществлять:

- все действия, доступные внутреннему нарушителю второй категории;
- модификацию ПО прикладного уровня системы АЛСР.

Внутренний нарушитель седьмой категории может располагать следующей информацией:

- любая информация, доступная внутреннему нарушителю второй категории;
- техническая документация на систему АЛСР;
- эксплуатационная документация на технические средства, используемые в АЛСР;
- алгоритм работы ПО прикладного уровня;
- любые фрагменты информации, обрабатываемой ПО прикладного уровня.

Внутренний нарушитель восьмой категории может осуществлять:

– реализацию конструктивно-встроенных программно-аппаратных закладок в технические средства, используемые на канальном и прикладном уровнях системы АЛСР.

Внутренний нарушитель восьмой категории может располагать следующей информацией:

– весь комплект документации на используемые в АЛСР технические средства;
– любые фрагменты информации, обрабатываемой в АЛСР.

В результате анализа разработанной модели нарушителя было установлено, что возможности внутренних нарушителей не зависят от наличия и использования радиоканала в системе АЛС. Следовательно, можно сделать вывод, что возможности внутренних нарушителей одинаковы для традиционных систем АЛС и систем АЛС на базе радиоканала. В связи с этим защита от угроз со стороны внутреннего нарушителя в АЛСР может быть реализована аналогично системам АЛС. Для нейтрализации данных угроз могут быть использованы способы, применяемые для обеспечения транспортной безопасности железнодорожных объектов [15], а также организационные и физические меры защиты.

Оставшиеся угрозы, влияющие на безопасность перевозочного процесса и связанные с применением в системе АЛС радиоканала, представлены в таблице 1.

Таблица 1 – Угрозы нарушения безопасности перевозочного процесса, связанные с использованием в АЛС радиоканала

№	Название	Параметр	Значение/Описание
1	Подмена базовой и/или абонентской радиостанции	Источник	Внешний нарушитель
		Уязвимость	Отсутствие процедуры взаимной аутентификации абонентов и базовой станции
		Способ реализации	Подавление сигнала базовой/абонентской радиостанции и поднятие на частоте ее радиообмена собственной радиостанции, передающей поддельную информацию
		Объект воздействия	Ответственная информация в радиоканале
		Деструктивное воздействие	Нарушение целостности ответственной информации
2	Отправка поддельной ответственной информации в радиоканал	Источник	Внешний нарушитель
		Уязвимость	Отсутствие контроля целостности передаваемой ответственной информации
		Способ реализации	Создание пакета данных с поддельной ответственной информацией и отправка его базовой/абонентской радиостанции
		Объект воздействия	Ответственная информация в радиоканале
		Деструктивное воздействие	Нарушение целостности ответственной информации
3	Повторная отправка ранее перехваченной в радиоканале	Источник	Внешний нарушитель
		Уязвимость	Отсутствие временной метки или порядкового номера пакета при передаче ответственной

	ответственной информации.		информации
		Способ реализации	Перехват легитимного сообщения, передаваемого по радиоканалу, и его повторная отправка участнику движения через некоторый период времени
		Объект воздействия	Ответственная информация в радиоканале
		Деструктивное воздействие	Нарушение целостности ответственной информации

Заключение

Таким образом, в результате анализа существующих угроз безопасности ответственной информации, передаваемой системой АЛСР, был определен перечень угроз, реализация которых из-за пределов контролируемой зоны может привести к нарушению безопасности перевозочного процесса. К ним относятся угрозы, связанные с нарушением целостности ответственной информации, передаваемой системой АЛСР, реализуемые со стороны внешнего нарушителя:

- подмена базовой и/или абонентской радиостанции;
- отправка поддельной ответственной информации в радиоканал;
- повторная отправка ранее перехваченной в радиоканале ответственной информации.

Приведенные в статье результаты работы могут использоваться при исследовании безопасности систем АЛСР, анализе существующих решений по защите ответственной информации, передаваемой АСЛР, или разработке новых способов защиты.

СПИСОК ЛИТЕРАТУРЫ:

1. О железнодорожном транспорте в Российской Федерации [Текст]: Федеральный закон от 10.01.2003 №17-ФЗ (ред. от 03.07.2016). Собрание законодательства Российской Федерации. – 2003. – № 2, (13 января). – статья 169.
2. Андерс Э. Системы автоматики и телемеханики на железных дорогах мира [Текст]: учебное пособие для вузов ж.-д транспорта Андерс Э., Берндт Т.; пер. с англ.; ред. Г. Теега, С. Власенко. – М.: Интекст, 2010. – 496 с. – ISBN 978-3-7771-0394-5.
3. ГОСТ Р 53431-2009. Автоматика и телемеханика железнодорожная. Термины и определения (с Изменением № 1) [Текст]. – Введ. 2011-01-01. – М.: Стандартинформ, 2010. – 23 с.
4. ОСТ 32.17-92. Безопасность железнодорожной автоматики и телемеханики. Основные понятия. Термины и определения [Текст]. – МПС России. – Введ. 1993-01-01. – СПб.: Типография ПИИТ, 1992. – 33 с.
5. Тильк И.Г. АЛС с использованием радиоканала [Текст] Тильк И.Г. Автоматика Связь Информатика. – 2010. – №7 – С. 7-9.
6. Шубинский И.Б. Немного о кибербезопасности [Текст] Шубинский И.Б., Макаров Б.А. Автоматика Связь Информатика. – 2014. – №8. С. 15-18.
7. Шубинский И.Б. О киберзащищенности информационных систем управления движением поездов [Текст] Шубинский И.Б., Макаров Б.А. Автоматика Связь Информатика. – 2014. – №11. – С. 6-12.
8. Макаров Б.А. Актуальность кибербезопасности на железнодорожном транспорте [Текст] Макаров Б.А. Техника железных дорог. – 2015. – №3 (31). – С. 19-24.
9. Гордейчик С.В. Кибербезопасность микропроцессорных систем ЖАТ [Текст] Гордейчик С.В. Автоматика Связь Информатика. – 2015. – №4. – С. 4-6.
10 Гордейчик С.В. Модель киберугроз МПЦ [Текст] Гордейчик С.В., Грицай Г.С., Баранов Д.С. Автоматика Связь Информатика. – 2015. – №7. – С. 18-20.
11. Konyavskiy V. The design of integrity monitoring and reliability verification system for critical information, transmitted in automatic train signaling system, based on DMR-RUS radio channel [Текст] Konyavskiy V., Epishkina A., Korotin A. Procedia Computer Science. – 2016. – Volume 88C. – P. 318-323.
12. Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка) [Электронный ресурс]: утверждена ФСТЭК России 15.02.2008. – Режим доступа: http://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-

dokumenty/379-bazovaya-model-ugroz-bezopasnosti-personalnykh-dannykh-pri-ikh-obrabotke-v-informatsionnykh-sistemakh-personalnykh-dannykh-vypiska-fstek-rossii-2008-god (дата обращения: 03.04.2017).
13. О системе технического обслуживания и ремонта локомотивов ОАО «РЖД» [Электронный ресурс]: Распоряжение ОАО «РЖД» от 17.01.2005 г. № 3р. – Режим доступа: http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=546128#0 (дата обращения: 03.04.2017).
14. Об утверждении правил технической эксплуатации железных дорог Российской Федерации [Электронный ресурс]: Приказ Министерства транспорта РФ от 21.12.2010 № 286. – Режим доступа: http://doc.rzd.ru/doc/public/ru?id=4051&layer_id= 5104&STRUCTURE_ID=704 (дата обращения: 03.04.2017).
15. О транспортной безопасности [Текст]: Федеральный закон от 09.02.2007 №16-ФЗ Собрание законодательства Российской Федерации. – 2007. – № 7. (12 февраля). – статья 837.

REFERENCES:

1. O zheleznodorozhnom transporte v Rossiyskoy Federatsii [Text]: Federalniy zakon ot 10.01.2003 №17-FZ (red. ot 03.07.2016) Sobraniye zakonodatelstva Rossiyskoy Federatsii. – 2003. – № 2. (13 yanvarya). – statya 169. (in Russian).
2. Anders E. Systems of automatics and telemechanics on the Railways of the world [Text]: textbook for universities.-d transportation Anders. E., Berndt. T.; lane. from English.; by G. Teega. S. Vlasenko. – M.: Intext. 2010. – 496 p. – ISBN 978-3-7771-0394-5. (in Russian).
3. GOST R 53431-2009. Avtomatika i telemekhanika zheleznodorozhnaya. Terminy i opredeleniya (s Izmeneniyem № 1) [Text]. – Vved. 2011-01-01. – M.: Standartinform. 2010. – 23 p. (in Russian).
4. OST 32.17-92. Bezopasnost' zheleznodorozhnoy avtomatiki i telemekhaniki. Osnovnyye ponyatiya. Terminy i opredeleniya [Text]. – MPS Rossii. – Vved. 1993-01-01. – SPb.: Tipografiya PIIT. 1992. – 33 p. (in Russian).
5. Tilk I. G. ALS using the radio [Text] Tilk I. G. Automation Communication Informatics. – 2010. – No. 7 – P. 7-9. (in Russian).
6. Shubinskiy I. B. a Little bit about cyber security [Text] Shubinskiy I. B., Makarov. B. A. Automation. Communication Informatics. – 2014. – No. 8. S. 15-18. (in Russian).
7. Shubinskiy I. B. On cybercasino information systems for train traffic control [Text] Shubinskiy I. B., Makarov. B. A. Automation. Communication Informatics. – 2014. – No. 11. – P. 6-12. (in Russian).
8. Makarov B. A. the Relevance of cyber security on railway transport [Text] Makarov B. A. the Moscow times. – 2015. – №3 (31). – P. 19-24. (in Russian).
9. S. V. Gordeychik Cybersecurity microprocessor based railway automatics and telemechanics systems [Text] S. V. Gordeychik Automation Communication Informatics. – 2015. – No. 4. – P. 4-6. (in Russian).
10 Gordeychik. S. V. Model of cyber threats MOC [Text] Gordeychik. S. V., Gritsai. G. S., Baranov D. S. Automation Communication Informatics. – 2015. – No. 7. – P. 18-20. (in Russian).
11. Konyavskiy V. The design of integrity monitoring and reliability verification system for critical information. transmitted in automatic train signaling system. based on DMR-RUS radio channel [Text] Konyavskiy V., Epishkina A., Korotin A. Procedia Computer Science. – 2016. – Volume 88C. – p. 318-323.
12. The basic model of threats to the security of personal data during their processing in personal data information systems (extract) [Electronic resource]: approved by the FSTEC of Russia 15.02.2008. – Access: http://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/379-bazovaya-model-ugroz-bezopasnosti-personalnykh-dannykh-pri-ikh-obrabotke-v-informatsionnykh-sistemakh-personalnykh-dannykh-vypiska-fstek-rossii-2008-god (data obrashcheniya: 03.04.2017). (in Russian).
13. O sisteme tekhnicheskogo obsluzhivaniya i remonta lokomotivov OAO «RZD» [Elektronnyy resurs]: Rasporyazheniye OAO «RZD» ot 17.01.2005 g. № 3r. – Access: http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=546128#0 (data obrashcheniya: 03.04.2017). (in Russian).
14. Ob utverzhdenii pravil tekhnicheskoy ekspluatatsii zheleznykh dorog Rossiyskoy Federatsii [Elektronnyy resurs]: Prikaz Ministerstva transporta RF ot 21.12.2010 № 286. – Access: http://doc.rzd.ru/doc/public/ru?id=4051&layer_id= 5104&STRUCTURE_ID=704 (data obrashcheniya: 03.04.2017). (in Russian).
15. O transportnoy bezopasnosti [Text]: Federal'nyy zakon ot 09.02.2007 №16-FZ Sobraniye zakonodatel'stva Rossiyskoy Federatsii. – 2007. – № 7. (12 fevralya). – statya 837. (in Russian).

Поступила в редакцию – 20 февраля июля 2017 г. Окончательный вариант – 21 мая 2017 г.
Received – February 20, 2017. The final version – May 21, 2017.

Андрей Н. Кузнецов, Дмитрий А. Вышемирский
Лаборатории ТВП, г. Москва, Нахимовский пр-т, дом 47,117418, Россия
e-mail: kuzz2000@mail.ru, ORCID iD 0000-0002-1782-6584
e-mail: v.dmitry-12@yandex.ru, ORCID iD 0000-0002-9015-7022

ОБ ОДНОМ ПОДХОДЕ К РЕШЕНИЮ ЗАДАЧИ ТОКЕНИЗАЦИИ ПРИ АНАЛИЗЕ БОЛЬШИХ МАССИВОВ ПОЛЬЗОВАТЕЛЬСКИХ ПАРОЛЕЙ

DOI: http://dx.doi.org/10.26583/bit.2017.2.06

Аннотация. Проведен анализ алгоритма выделения в паролях пользователей слов на естественных языках (токенизации), предложенного ранее в работе R. Veras et al. [1] (RV-алгоритм). Выявлены основные недостатки данного подхода. Предложен новый алгоритм токенизации (RGramToken) на основе частотных словарей слов, биграмм и триграмм, позволяющий лучше учесть информацию о вероятности использования слов и словосочетаний в естественном языке. Подготовлены тестовые выборки фраз с учетом возможных искажений типа «вставка» между словами естественного языка. Проведен сравнительный анализ результатов работы на тестовых и реальных выборках алгоритма RGramToken с RV-алгоритмом. Показано более высокое качество токенизации предложенным в работе алгоритмом на искаженных фразах, а также устойчивость к качеству используемых словарей.

Ключевые слова: пароли, утечки, токенизация, разбиение на слова.

Для цитирования. КУЗНЕЦОВ, Андрей Н.; ВЫШЕМИРСКИЙ, Дмитрий А. Об одном подходе к решению задачи токенизации при анализе больших массивов пользовательских паролей. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 50-60, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/105>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.06.

Andrey N. Kuznetsov, Dmitry A. Vyshemirsky
TVP laboratories, Moscow, Nakhimovsky PR-t, 47, 117418, Russia
e-mail: kuzz2000@mail.ru, ORCID iD 0000-0002-1782-6584
e-mail: v.dmitry-12@yandex.ru, ORCID iD 0000-0002-9015-7022

One approach to solving tokenization problem for analysis of large-scale collections of user-defined passwords

DOI: http://dx.doi.org/10.26583/bit.2017.2.06

Abstract. This paper performs an analysis of the algorithm of password tokenization introduced by R. Veras et al. [1]. We show main limitations of this approach and propose a new tokenization algorithm - RGramToken, based on frequency dictionaries of English words, bigrams and trigrams. Our approach allows better utilization of information about probabilitiy distribution of words and word combinations in a natural language. The results of comparison analysis of these two algorithms on specially prepared tests with warped phrases demonstrate higher efficiency of RGramToken and its robustness on low quality dictionaries.

Keywords: passwords, leaks, tokenization, words isolation.

For citation. KUZNETSOV, Andrey N.; VYSHEMIRSKY, Dmitry. A. One Approach to Solving Tokenization Problem for Analysis of Large-Scale Collections of User-Defined Passwords. IT Security (Russia), [S.l.], v. 24, n. 2, p. 50-60, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/105>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.06.

Введение

Парольная защита на сегодняшний день является одним из наиболее распространенных способов решения задач аутентификации пользователей и обеспечения конфиденциальности пользовательских данных. Распространенность парольной защиты

как элемента криптографических систем практически не снижается, несмотря на многочисленные компрометации, получившие широкую огласку в последние годы: утечки аутентификационных данных миллионов пользователей в сеть Интернет, а также некоторые результаты анализа таких утечек. На примере многих сайтов показано, что знание всего 10 наиболее популярных паролей позволяет получить доступ к данным 5-10% пользователей из их многомиллионной аудитории [2].

Статистики распределения длин паролей, а также структуры используемых в них алфавитов, полученные в результате анализа таких баз реальных паролей, позволяют оценить эффективность тотального опробования. Однако, подсчет теоретико-информационных характеристик парольных массивов, в особенности значения энтропии текста [3], показал, что построение более сложных моделей парольных источников, учитывающих закономерности в выборе паролей и зависимости между символами и группами символов внутри пароля, позволяет существенно повысить эффективность атак, с одной стороны, и сформулировать требования к выбираемым паролям, повышающие их безопасность, с другой [4].

Появились работы, в которых предлагаются различные стохастические подходы к моделированию парольных источников: вероятностные контекстно-независимые грамматики [5-7], марковские модели [8-10], другие оригинальные модели (например, [3]). Практически все известные авторам подходы, включая перечисленные, основаны на статистических зависимостях между символами или группами символов из различных алфавитов. При этом в практике автоматизированного анализа текстов на естественных языках (*Natural Language Processing – NLP*) стандартом де-факто на сегодня являются модели, учитывающие зависимости между словами, которые гораздо лучше отражают реальные свойства текстов.

Для построения модели пароля, в основе которой будут лежать слова естественного языка, необходимо иметь адекватную обучающую выборку, то есть для расчета статистик слов необходимо иметь массив реальных паролей с разметкой по словам. Это, в свою очередь, означает, что необходимо решить так называемую задачу *токенизации* – разделения текста на слова.

Для текстов в традиционном понимании (художественных, публицистических, деловых и т.п.) подходы к решению данной задачи хорошо известны [11, 12] и достаточно эффективны, однако специфика построения паролей – умышленные искажения слов и вставка семантически необусловленных аффиксов – делает эти методы практически неприменимыми. Для массивов паролей необходим новый подход, позволяющий учесть особенности их построения.

Первую известную авторам попытку разработать алгоритм токенизации для паролей предприняли в 2013 году M. Jakobsson и M. Dhiman [13]. В 2014 году команда исследователей из канадского университета Онтарио (R. Veras et al.) предложила модификацию данного алгоритма [1], сочтя первоначальную версию недостаточно эффективной.

Попытка реализовать алгоритм 2014 года на практике и проверить на доступных базах реальных паролей показала неудовлетворительный результат. В настоящей работе делается попытка проанализировать причины низкой эффективности подхода R. Veras и предлагается новый алгоритм выделения слов естественного языка в пользовательских паролях, а также проводится сравнительный анализ эффективности предложенного подхода с решением группы из Онтарио.

Алгоритм R. Veras (RV-алгоритм)

В работе [1] алгоритм токенизации описан достаточно подробно, здесь мы кратко изложим его основные принципы.

В основе предложенного подхода лежит набор словарей имен собственных – имена, фамилии, страны, города и названия месяцев, а также корпус современного английского языка с частотами встречаемости отдельных слов, их биграмм и триграмм (*Contemporary*

Corpus of American English – COCA). Поскольку корпус COCA распространяется на коммерческой основе (бесплатно доступна лишь ограниченная онлайн версия), мы в своих исследованиях использовали свободно распространяемый аналог – корпус *ENCOW* (*English web corpus by COW*) [14].

Пароль представляется как последовательность слов и вставок, не являющихся словами естественного языка. Авторы [1], аналогично подходу, предложенному в [13], строят для каждого пароля все возможные варианты сегментации и вычисляют степень покрытия пароля словарными словами. Вариант с максимальным значением покрытия считается правильным. Если таких вариантов оказывается несколько, для каждого из них вычисляется значение правдоподобия по статистике встречаемости слов и их сочетаний (*N-gram score*). В работе рассматриваются значения $N \in \{1,2,3\}$.

Алгоритм вычисления значения правдоподобия (*score*) изложен авторами в виде псевдокода (рисунок 1).

Algorithm 2 Recursively calculate the N-gram score of a segmentation

```
1: procedure BESTNGRAMSCORE(C)
2:     score ← 0
3:     l ← LENGTH(C)
4:
5:     if l = 1 then
6:         score ← UNIGRAMPROBABILITY(C)
7:     else if l = 2 then
8:         score ← BIGRAMPROBABILITY(C)
9:     else if l = 3 then
10:        score ← TRIGRAMPROBABILITY(C)
11:    end if
12:
13:    if score = 0 then
14:        for i ← 1, 3 do
15:            a ← BESTNGRAMSCORE(C[: i])
16:            b ← BESTNGRAMSCORE(C[i :])
17:            tempScore ← a * b
18:            if tempScore > score then
19:                score ← tempScore
20:            end if
21:        end for
22:    end if
23: end procedure
```

Рис.1 Алгоритм подсчета правдоподобия

По крайней мере два решения в предлагаемом подходе могут приводить к ошибкам в результатах токенизации:

1. выбор кандидатов по максимальному значению покрытия;
2. способ вычисления величины правдоподобия, когда безальтернативно используется вероятность r-грамм с максимальным значением r.

Разберем эти ситуации на примерах.

Максимизация величины покрытия приводит на практике к тому, что алгоритм предпочтет несколько несвязанных друг с другом низковероятных слов длинному высокочастотному слову, если оно будет хоть на символ короче суммы длин слов в первом варианте.

Проиллюстрировать сказанное можно примером токенизации реального пароля «threeflattenedf», который, очевидно, получен из фразы «three flattened» путем вставки концевого символа «f». Наибольшим покрытием при этом обладает разбиение «three flatten edf» без вставок. Все слова полученной фразы присутствуют как в корпусе COCA, так в корпусе ENCOW. Вероятности r-грамм для вычисления вероятности данного и оригинального разбиений по корпусу ENCOW приведены в таблице 1.

Таблица 1. Частоты и вероятности r-грамм, входящих в пароль «threeflattenedf»

***r*-грамма**	**Частота**	**Вероятность**

three	$3{,}96*10^{6}$	$5{,}64*10^{-4}$
flattened	$2{,}09*10^{4}$	$2{,}98*10^{-6}$
flatten	$9{,}86*10^{3}$	$1{,}40*10^{-6}$
edf	187	$2{,}66*10^{-8}$
three flattened	12	$2{,}07*10^{-9}$
three flatten	1	$1{,}73*10^{-10}$
flatten edf	0	0
three flatten edf	0	0

Значение правдоподобия разбиения «three flatten edf» составляет p(«three flatten»)*p(«edf») = $4{,}68*10^{-18}$, в то время как для разбиения «three flattened» оно составляет p(«three flattened»)=$2{,}07*10^{-9}$. То есть при вероятности вставки буквы «f» в конце фразы выше $2{,}26*10^{-9}$ вариант разбиения исходной цепочки с вставкой в конце окажется более предпочтительным, но будет отбракован по величине покрытия.

В случае второго замечания если для рассматриваемого варианта разбиения в частотном корпусе находится триграмма (биграмма) с ненулевой вероятностью, то при вычислении правдоподобия всего варианта будет использовано это значение вероятности, даже если с точки зрения максимизации правдоподобия выгоднее рассматривать эту последовательность как набор r-грамм меньшего порядка.

Примером в данном случае может служить цепочка «aalex», при токенизации которой максимальное покрытие достигается для двух вариантов разбиения: «a a lex» и «a alex». Вероятности, необходимые для вычисления правдоподобия полученных разбиений, приведены в таблице 2. В таблице 3 вычислены правдоподобия разбиений при разных комбинациях одограмм, биграмм и триграмм. Все значения условных часот взяты из корпуса ENCOW.

Таблица 2. Частоты и вероятности *r*-грамм, входящих в пароль «aalex»

***r*-грамма**	**Частота**	**Вероятность**
a	$1{,}79*10^{8}$	$2{,}5*10^{-2}$
lex	2’789	$3{,}97*10^{-7}$
alex	3’787	$5{,}39*10^{-7}$
a a	22’388	$3{,}86*10^{-6}$
a lex	94	$1{,}62*10^{-8}$
a alex	2	$3{,}45*10^{-10}$
a a lex	0	0

Таблица 3. Правдоподобия всех возможных разбиений цепочки символов «aalex» согласно RV-алгоритму

Разбиение	**Комбинация *r*-грамм**	**Правдоподобие**
a a lex	p(«a a lex»)	0
	p(«a a») * p(«lex»)	$1{,}53*10^{-13}$
	p(«a») * p(«a lex»)	**$4{,}05*10^{-10}$**
a alex	p(«a alex»)	$3{,}45*10^{-10}$
Игнорируемые алгоритмом комбинации *r*-грамм		
a a lex	p(«a») * p(«a») * p(«lex»)	$2{,}48*10^{-10}$
a alex	**p(«a») * p(«alex»)**	**$1{,}35*10^{-8}$**

Из таблицы 3 видно, что исходный алгоритм вычисления правдоподобия результатов токенизации предпочел бы вариант «a a lex», который с мнемонической точки зрения менее привлекателен, чем «a alex». В случае разбиения биграммы (a, alex) на две одограммы, более правдоподобным оказался бы вариант «a alex».

Новый алгоритм токенизации (RGramToken)

Исходя из анализа достоинств и недостатков рассмотренного RV-алгоритма, предлагается новый алгоритм токенизации паролей пользователей. Ниже приводится его формальное описание.

Пусть задан словарь $S = \{s_1, \dots, s_n\}$, $n \in \mathbb{N}$, где s – слово естественного языка, состоящее из символов некоторого алфавита $\boldsymbol{A}$. Пусть заданы также статистики распределения частот r-грамм слов словаря S в некой обучающей выборке $\nu_r = \nu(w_1, \dots, w_r)$ для всех $w \in S$, $r = 1 \div 3$, где $\nu(w_1, \dots, w_r)$ – частота встречаемости r-граммы $w_1, \dots, w_r$.

Будем рассматривать биграммы и триграммы слов с ненулевыми частотами как самостоятельные слова с длиной равной сумме длин входящих в них слов, т.е. построим словарь $S' = S \cup S_2 \cup S_3$, где S_2 и S_3 – словари биграмм и триграмм с ненулевыми частотами соответственно. Тогда мы имеем общий вектор частот $\nu' = \nu(w_i')$, $w_i' \in S'$.

Множество всех возможных несловарных вставок в паролях обозначим через $V = \{\mathrm{v}_i\} \cup \{\emptyset\}$, где v_i – цепочки символов алфавита $\boldsymbol{B}$, в общем случае равного либо включающего $\boldsymbol{A}$, длиной не более максимальной длины пароля l_{max}, таких, что $\mathrm{v}_i \notin S$.

Необходимо представить произвольный пароль $D = (d_1, \dots, d_l)$, $l \in \mathbb{N}$, $d \in \boldsymbol{B}$ в виде последовательности слов словаря S и вставок несловарных последовательностей из множества V оптимальным способом в соответствии с описанным ниже критерием.

На первом этапе строятся все возможные варианты разбиения пароля D. Для всех $i \in [1, l]$, для которых $(d_1, \dots d_{i-1}) = \mathrm{v}_{j_i} \in V$, выбрать все возможные словарные слова $(d_i, \dots, d_{k_i}) = s_{m_i} \in S$, начинающиеся с i-го символа (при $i = 1$ вставка v_{j_i} будет нулевой длины). В результате будут получены цепочки вида $\{\mathrm{v}_{j_i}, s_{m_i}, (d_{k_i+1}, \dots d_l)\}$, для каждой из которых необходимо произвести разбиение нового пароля $D_{k_i} = (d_{k_i+1}, \dots, d_l)$ и конкатенировать результат с уже полученным началом $\{\mathrm{v}_{j_i}, s_{m_i}\}$.

Все полученные последовательности вида $\{\mathrm{v}_1, s_{11}, \dots s_{1k_1}, \mathrm{v}_2, s_{21}, \dots s_{2k_2}, \dots \mathrm{v}_p\}$, где только v_1 и v_p могут быть пустыми, а последовательности слов $s_{i1}, \dots s_{ik_i}$ при любых значениях i не содержат внутри себя цепочек из множества V, являются результатами разбиения пароля D.

Далее на втором этапе среди множества вариантов необходимо выбрать разбиение, на котором достигается максимум функции правдоподобия. Для этого вычислить величину правдоподобия для каждого из разбиений на основе общего вектора частот $\nu' = \nu(w_i')$, $w_i' \in S'$. Предварительно необходимо вычислить значения относительных частот (вероятностей) по формуле:

$$p(w_i') = \frac{\nu(w_i')}{\sum_{w_i' \in S'} \nu(w_i')}$$

Значение правдоподобия варианта разбиения $\{\mathrm{v}_1, s_{11}, \dots s_{1k_1}, \mathrm{v}_2, s_{21}, \dots s_{2k_2}, \dots \mathrm{v}_p\}$, полученного на первом этапе, вычисляется по формуле

$$L(\mathrm{v}_1, s_{11}, \dots s_{1k_1}, \mathrm{v}_2, s_{21}, \dots s_{2k_2}, \dots \mathrm{v}_p) = L\left(s_{11}, \dots, s_{(p-1)k_{p-1}}\right) \prod_{i=1}^{p} L(\mathrm{v}_i).$$

Для этого цепочка словарных слов $\left(s_{11}, \dots, s_{(p-1)k_{p-1}}\right)$ всеми возможными способами представляется как последовательность $\overline{w'}_j = \left(w'_{i\,j}\right)$ элементов множества S'. Далее из множества $\{\overline{w'}_j\}$ возможных последовательностей выбирается вариант с максимальным значением правдоподобия, то есть $L\left(s_{11}, \dots, s_{(p-1)k_{p-1}}\right) = \max_j L(\overline{w'}_j)$, где $L(\overline{w'}_j) = \prod p(w'_{i_j})$.

Правдоподобие вставки для простоты предлагается вычислять по формуле $L(\mathrm{v}_i) = p(\mathrm{v}_0)^{len(\mathrm{v}_i)}$, где $p(\mathrm{v}_0)$ – вероятность вставки одного символа, $len(\mathrm{v}_i)$ – длина вставки v_i в символах. Здесь мы исходим из предположения, что короткие вставки между словами

естественного языка более вероятны, чем длинные. Значение $p(\mathrm{v}_0)$ есть параметр алгоритма и может влиять на результаты токенизации. В дальнейшем можно усложнить алгоритм, введя в него более сложные модели несловарных вставок в паролях. Для пустых вставок $L(\{\emptyset\}) = 1$.

Выбор наилучшего варианта разбиения пароля D производится по критерию максимума правдоподобия, то есть как $argmax\left(L(\mathrm{v}_1, s_{11}, \dots s_{1k_1}, \mathrm{v}_2, s_{21}, \dots s_{2k_2}, \dots \mathrm{v}_p)\right)$.

Результаты экспериментов

Как было сказано выше, для построения словаря S' и соответствующего массива частот ν' в настоящей работе использовался корпус ENCOW версии мая 2015 года. В качестве слов в данной выборке участвуют не только английские слова, но и различные знаки (кавычки, знаки препинания, апострофы, числа). Слова записаны с учетом регистра.

В ходе предварительного этапа из корпуса были исключены все r-граммы, содержащие символы не из английского алфавита, при этом апострофы и кавычки были предварительно удалены, а буквы приведены к нижнему регистру. Совпадающие r-граммы объединялись, а их частоты складывались.

Полученный в результате корпус оказался сильно зашумлен искаженными словами с низкими частотами, поэтому из него были исключены слова с частотой ниже выбранного порога v_b. Для оценки влияния значения порога частоты на результат алгоритма были подготовлены корпуса с использованием значений $v_b = 5$ (962 067 слов) и $v_b = 50$ (243 077 слов). Такое усечение словаря исключает из возможных вариантов разбиения редкие фамилии, географические названия и т.п. Этого можно избежать путем добавления в результирующий корпус слов из специализированных словарей имен, названий и т.д., как было сделано в работе [1]. Значение частоты для добавляемых слов должно находиться в интервале $1 \leq \nu \leq v_b$. В рамках данного исследования такая процедура не проводилась.

Для проведения сравнительного анализа алгоритма RGramToken, предложенного в настоящей работе, и RV-алгоритма, проведен ряд экспериментов на тестовых примерах. В качестве тестовой выборки выступал специальным образом подготовленный словарь фраз, построенный на основе художественного произведения «Crashlander» Ларри Нивена на английском языке. Из текста были удалены все символы, кроме букв английского алфавита и пробела, символы переведены в нижний регистр, после чего удалены слова, не входящие в исходный словарь. Далее весь текст был разделен на N фраз так, чтобы каждая i-я фраза содержала k_i слов, сумма длин первых $k-1$ слов $l_{k_i-1} < 12$, а всех слов фразы $12 \leq l_{k_i} \leq 20$ символов без учета пробелов. В итоге размер полученного словаря составил $N = 16646$ фраз.

Далее на основе базового тестового словаря было построено три словаря с искаженными фразами. Каждая фраза искажалась путем вставки случайных несловарных последовательностей символов длины $1 \leq m \leq 5$, значение m выбиралось случайно и равновероятно. Для оценки устойчивости результатов вставки делались тремя способами: в конец фразы, в середину (в случайную позицию между двумя словами), а также одновременно в середину и конец фразы.

После применения исследуемых алгоритмов к каждой фразе тестовых словарей подсчитывались число успешных разбиений и их доля от общей мощности словаря. Успешным считалось точное восстановление исходного состава фразы с указанием вставок.

Результаты токенизации исходных и искаженных фраз RV-алгоритмом и алгоритмом RGramToken приведены в таблице 4. Эксперименты проводились с использованием корпуса ENCOW с отсечением низкочастотных слов по порогам $v_b = 5$ и $v_b = 50$, алфавиты слов и вставок $|\boldsymbol{A}| = |\boldsymbol{B}| = 26$, вероятность вставки одного символа $p(\mathrm{v}_0) = 1/26$.

Таблица 4. Результаты токенизации на тестовых выборках

Тестовый словарь	**RV-алгоритм**		**RGramToken**	
	$v_b = 5$	$v_b = 50$	$v_b = 5$	$v_b = 50$
Без вставок	16085 (96,6%)	15602 (93,7%)	12284 (73,8%)	12250 (73,6%)
В конец	3974 (23,9%)	5016 (30,1%)	8115 (48,8%)	8116 (48,8%)
В середину	1821 (10,9%)	2885 (17,3%)	11760 (70,7%)	11647 (70,0%)
В конец и в середину	512 (3,1%)	970 (5,8%)	7803 (46,9%)	7759 (46,6%)

Приведенные результаты иллюстрируют неудовлетворительную работу RV-алгоритма на искаженных фразах. Скорее всего, авторы использовали подходы, разработанные для токенизации текстов с опущенными пробельными символами, что не позволило показать хорошие результаты на зашумленных фразах, которые в большей мере соответствуют реальным пользовательским паролям. Особенно хорошо это видно на словаре с двумя искажениями, где RV-алгоритм уступил более чем в 6 раз. Алгоритм RGramToken гораздо успешнее справился с искаженными примерами, хотя и уступил в задаче токенизации неискаженных фраз.

Тот факт, что вставки в середину существенно лучше обрабатываются алгоритмом RGramToken, чем вставки в конец, наглядно демонстрирует выигрыш от более полного использования информации о частотах биграмм и триграмм, что позволяет выбрать более правильные грамматические формы словарных слов.

Эксперименты с различными значениями порога отсечения по частоте встречаемости слов показали, что RV-алгоритм, в отличие от алгоритма RGramToken, чувствителен к качеству используемого словаря, поэтому использование коммерческого корпуса авторами работы [1] представляется оправданным.

В заключение приведем некоторые результаты эксперимента по токенизации на массиве реальных паролей. В качестве источника таких паролей использован известный массив аккаунтов пользователей сайта Rockyou.com, который в 2009 году в результате утечки попал в сеть Интернет [15]. Предварительно все буквы английского алфавита в паролях переведены в нижний регистр. Результаты работы RV-алгоритма и алгоритма RGramToken с порогом отсечения $v_b = 50$ приведены в таблице 5. Вставки обрамлены с помощью пар символов $> слева и <$ справа (т.е. вставка «1» представляется в виде $>1<$), слова и вставки отделены друг от друга пробелами.

Таблица 5. Примеры разбиения реальных паролей

Пароль	**Разбиение**	
	RV-алгоритм	**RGramToken**
«bedroomeyez»	bed roome yez	bedroom eye $>z<$
«crisine»	chris ine	chris in $>e<$
«girl090pink»	girl $>090<$ pink	girl $>090p<$ in $>k<$
«lovearizona»	love arizona	love arizona
«gopalloveforu»	gopal love foru	go pal love for $>u<$
«martinsac1»	martin sac $>1<$	martins a $>c1<$
«samanthaco»	sama nt hac $>o<$	samantha $>co<$
«silvestrek»	silves trek	silvestre $>k<$
«whitethouse»	white thouse	white $>t<$ house

Объективно определить лучший из двух рассмотренных алгоритмов на множестве реальных паролей не представляется возможным из-за отсутствия исходных фраз, положенных в основу пароля пользователями, поэтому здесь приводятся результаты без оценок качества.

Заключение

Алгоритм RGramToken, предложенный в данной работе для выделения слов естественного языка в реальных пользовательских паролях, позволяет эффективно производить токенизацию на больших массивах парольной информации, доступных для анализа в сети Интернет. Эффективность предложенного метода показана в сравнительных экспериментах с лучшим из известных авторам алгоритмов, описанным в работе R. Veras et al. [1].

Основными преимуществами нового алгоритма является устойчивость к умышленным искажениям исходных фраз путем добавления различных несловарных вставок, что является обычной практикой при выборе пароля. Кроме того, алгоритм RGramToken более эффективно использует информацию о частоте встречаемости слов и словосочетаний, что делает его менее чувствительным к качеству используемого словаря.

Результаты токенизации реальных парольных массивов могут быть использованы в дальнейшем для проведения новых исследований грамматического и семантического состава пользовательских паролей, строить адекватные лингвистические и теоретико-вероятностные модели парольных источников.

Последующие исследования в области выделения слов естественного языка в реальных пользовательских паролях могут быть связаны с изучением природы несловарных вставок и построением их математической модели. Важным для практического применения предложенного алгоритма представляется также формирование рабочего словаря, содержащего качественный частотный корпус и специальные словари имен собственных, известных аббревиатур и сокращений.

СПИСОК ЛИТЕРАТУРЫ:

1. Veras R., Collins C., Thorpe J. On the Semantic Patterns of Passwords and their Security Impact. NDSS. – 2014.

2. Wang D., Jian G., Wang P. Zipf's Law in Passwords. IACR Cryptology ePrint Archive. – 2014. – Т. 2014. – С. 631.

3. Тюрин К.А., Сёмин Р.В. Анализ стойкости парольных фраз на основе информационной энтропии. Известия Южного федерального университета. Технические науки. – 2015. - №5 (166). С. 18-27.

4. Марков Г.А. К вопросу об определении стойкости парольных систем. Сборник трудов Третьей всероссийской НТК «Безопасные информационные технологии». – М.: НИИ РЛ МГТУ им. Н.Э. Баумана, 2012. С. 21-23.

5. Weir M., Aggarwal S., De Medeiros B., Glodek B. Password cracking using probabilistic context-free grammars. 2009 IEEE Symposium on Security and Privacy. – IEEE, 2009. С. 391-405.

6. Houshmad S., Aggarwal S., Flood R. Next Gen PCFG Password Cracking. IEEE, 2015. С. 1776-1791.

7. Yazdi Sh. Probabilistic Context-Free Grammar Based Password Cracking: Attack, Defence and Application. FSU Libraries, 2015.

8. Van Heerden R.P., Vorster J.S. Using Markov Models to crack passwords. The 3rd International Conference on Information Warfare and Security: Peter Kiewit Institute, University of Nebraska, Omaha, USA. – 2008. С. 24-25.

9. Ma J., Yang W., Luo M., Li N. A Study of Probabilistic Password Models. 2014 IEEE Symposium on Security and Privacy. – IEEE, 2014. С. 689-704.

10. Duermuth M., Angelstorf F., Castellucia C., Perito D., Chaabane A. OMEN: Faster Password Guessing Using an Ordered Markov Enumerator. International Symposium on Engeneering Secure Software and Systems, 2015. С. 119-132.

11. Grefenstette G., Tapanainen P. What is a word, What is a sentence? Problems of Tokenization. Proceedings of the 3rd Conference on Computational Lexicography and Text Research, COMPLEX'94. – 1994. С. 79-87.

12. Jurish B., Wurzner K.-M. Word and Sentence Tokenization with Hidden Markov Models. JLCL. – 2013. – Т.28, №2. С. 61-83.

13. Jakobsson M., Dhiman M. The benefits of understanding passwords. Mobile Authentication, ser. Springer Briefs in Computer Science. – Springer New York, 2013. С. 5-24.

14. ENCOW14. [Электронный ресурс.] – URL: http://corporafromtheweb.org/encow14/ (дата обращения 30.05.2016).

15. RockYou. [Электронный ресурс.] – URL: http://wiki.skullsecurity.org/Passwords (дата обращения 30.05.2016).

REFERENCES:

[1]Veras R., Collins C., Thorpe J. On the Semantic Patterns of Passwords and their Security Impact. NDSS. – 2014.

[2]Wang D., Jian G., Wang P. Zipf's Law in Passwords. IACR Cryptology ePrint Archive. – 2014. – T. 2014. – p. 631.

[3]Turin K.A., Semin R.V. Analysis of Passphrases Resistance Based on Information Entropy. Izvestiya SFedU. Engeneering Sciences. – 2015. - №5 (166). pp. 18-27. (in Russian)

[4]Markov G.A. About password systems strength counting. Sbornik trudov Tret'ey vserossiyskoy NTK «Bezopasnye informatsionnye tehnologii». – Moscow. Bauman Moscow State Technical University, 2012. pp. 21-23. (in Russian)

[5] Weir M., Aggarwal S., De Medeiros B., Glodek B. Password cracking using probabilistic context-free grammars. 2009 IEEE Symposium on Security and Privacy. – IEEE, 2009. pp. 391-405.

[6] Houshmad S., Aggarwal S., Flood R. Next Gen PCFG Password Cracking. IEEE, 2015. pp. 1776-1791.

[7] Yazdi Sh. Probabilistic Context-Free Grammar Based Password Cracking: Attack, Defence and Application. FSU Libraries, 2015.

[8] Van Heerden R.P., Vorster J.S. Using Markov Models to crack passwords. The 3rd International Conference on Information Warfare and Security: Peter Kiewit Institute, University of Nebraska, Omaha, USA. – 2008. pp. 24-25.

[9] Ma J., Yang W., Luo M., Li N. A Study of Probabilistic Password Models. 2014 IEEE Symposium on Security and Privacy. – IEEE, 2014. pp. 689-704.

[10]Duermuth M., Angelstorf F., Castellucia C., Perito D., Chaabane A. OMEN: Faster Password Guessing Using an Ordered Markov Enumerator. International Symposium on Engeneering Secure Software and Systems, 2015. pp. 119-132.

[11]Grefenstette G., Tapanainen P. What is a word, What is a sentence? Problems of Tokenization. Proceedings of the 3rd Conference on Computational Lexicography and Text Research, COMPLEX'94. – 1994. pp. 79-87.

[12]Jurish B., Wurzner K.-M. Word and Sentence Tokenization with Hidden Markov Models. JLCL. – 2013. – T.28, №2. pp. 61-83.

[13] Jakobsson M., Dhiman M. The benefits of understanding passwords. Mobile Authentication, ser. Springer Briefs in Computer Science. – Springer New York, 2013. pp. 5-24.

[14] ENCOW14. Available at: http://corporafromtheweb.org/encow14/ (accessed 30.05.2016).

[15] RockYou. Available at: http://wiki.skullsecurity.org/Passwords (accessed 30.05.2016).

Поступила в редакцию – 20 февраля июля 2017 г. Окончательный вариант – 21 мая 2017 г.
Received – February 20, 2017. The final version – May 21, 2017.

Владимир А. Минаев, Екатерина В. Вайц, Юлия В. Грачёва
Московский государственный технический университет им. Н.Э. Баумана,
2-я Бауманская ул., 5, 105005, Россия
e-mail: m1va@yandex.ru, ORCID iD 0000-0002-4356-301X
e-mail: vaitcev@yandex.ru, ORCID iD 0000-0002-4629-6252
e-mail: uvgracheva@gmail.com, ORCID iD 0000-0003-0884-5617

ИМИТАЦИОННЫЕ ЭКСПЕРИМЕНТЫ С МОДЕЛЬЮ ИНФОРМАЦИОННО-ПСИХОЛОГИЧЕСКИХ ВОЗДЕЙСТВИЙ НА МАССОВОЕ СОЗНАНИЕ

DOI: http://dx.doi.org/10.26583/bit.2017.2.07

Аннотация. В статье рассматривается проблема исследования динамики информационно-психологических воздействий на массовое сознание, возможность их прогнозирования и управления, что является одним из важнейших аспектов обеспечения информационно-психологической безопасности общества и его граждан. Для исследования динамики информационно-психологических воздействий на массовое сознание в статье предложен метод системно-динамического моделирования, обоснованный и реализованный на моделях сложных социально-экономических явлений Дж. Форрестером в 1950-х годах. Применение данного метода к решению проблем обеспечения информационной безопасности исследуется различными зарубежными научными коллективами. Метод системно-динамического моделирования позволяет отобразить и исследовать множество факторов, критично влияющих на процессы информационно-психологических воздействий на массовое сознание. В предложенной авторами статьи модели учтены следующие факторы: вероятность «увлечения» идеями, заложенными в контент информационно-психологических воздействий, при межличностном контакте и в результате воздействия средств массовой информации, показатель массовости и регулярности средств массовой информации, пропагандирующих идею информационно-психологических воздействий, вероятность забывания идеи, заложенной в контент информационно-психологических воздействий, вероятность коммуникации на тему, заложенную в контент информационно-психологических воздействий, вероятность «увлечения» идеями информационно-психологических воздействий при одном коммуникационном контакте, среднее количество ознакомлений с одним сообщением в средствах массовой информации, вероятность «увлечения» новой идеей из контента информационно-психологических воздействий после ознакомления с сообщением в средствах массовой информации, количество контактов человека, увлеченного идеями информационно-психологических воздействий, в день. Математическим аппаратом, на котором базируются системно-динамические модели, являются системы дифференциальных уравнений. Авторами статьи построена имитационная системно-динамическая модель, позволяющая выявлять взаимосвязи между факторами, влияющими на динамику информационно-психологических воздействий на массовое сознание, и выполнена ее реализация в программной среде Anylogic. Данная среда является одной и самых перспективных сред имитационного моделирования. Приведены временные графики, отражающие динамику количества лиц, способных принять идею информационно-психологических воздействий, принявших идею информационно-психологических воздействий, а также графики, отражающие распределение во времени количества лиц, принявших идею информационно-психологических воздействий при определенных начальных условиях. Проведены имитационные эксперименты, связанные с варьированием параметров модели, в которых исследуется зависимость темпа увеличения количества лиц, принявших идею информационно-психологических воздействий, от частоты контактов лиц, разделяющих данные идеи.

Ключевые слова: системно-динамическое моделирование, имитационное моделирование, имитационные эксперименты, информационно-психологические воздействия, массовое сознание.

Для цитирования. МИНАЕВ, Владимир А.; ВАЙЦ, Екатерина В.; ГРАЧЁВА, Юлия В. Имитационные эксперименты с моделью информационно-психологических воздействий на массовое сознание. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 61-71, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/106>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.07.

Vladimir A. Minaev, Ekaterina V. Vaitc, Yulia V. Gracheva
Bauman Moscow State Technical University,
2nd Baumanskaya street, 5, 105005, Russia,
e-mail: m1va@yandex.ru, ORCID iD 0000-0002-4356-301X
e-mail: vaitcev@yandex.ru, ORCID iD 0000-0002-4629-6252
e-mail: uvgracheva@gmail.com, ORCID iD 0000-0003-0884-5617

Simulation Experiments with the Model of Information-Psychological Influences on Mass Consciousness

DOI: http://dx.doi.org/10.26583/bit.2017.2.07

Abstract. The article deals with the problem of researching the dynamics of information and psychological influences on mass consciousness, the possibility of their forecasting and management, which is one of the most important aspects of ensuring the information and psychological security of society and its citizens. To research the dynamics of information and psychological influence on mass consciousness, the article suggests a method of system-dynamic modeling, grounded and implemented on models of complex socio-economic phenomena by J. Forrester in the 1950s. The application of this method to solving the problems of information security is investigated by various foreign scientific teams. The method of system-dynamic modeling allows to display and investigate many factors that critically affect the processes of information and psychological influences on mass consciousness. The following factors are taken into account in the model proposed by the authors of the model: the probability of "enthusiasm" of ideas embedded in the content of information and psychological influences, with interpersonal contact and as a result of the influence of the mass media, the massiveness and regularity of the mass media propagandizing the idea of information and psychological influences, the probability of forgetting the idea, embedded in the content of information and psychological influences, the probability of communication on the topic, embedded in the content of information and psychological influences, the probability of "enthusiasm" ideas of information and psychological influences in one communication contact, the average number of acquaintances with one message in the media, the probability of "enthusiasm" of a new idea from the content of information and psychological influences after reading the information in the media, the number of contacts a person, keen on ideas of information and psychological influences, per day. The mathematical apparatus on which system-dynamic models are based are systems of differential equations. The authors of the article constructed an imitation model that makes it possible to identify the interrelationships between factors that effect the dynamics of information and psychological influences on mass consciousness, and its implementation has been carried out in the Anylogic software environment. This environment is one of the most promising environments for simulation. Temporal graphs reflecting the dynamics of the number of persons capable of adopting the idea of information and psychological influence, the number of persons who adopted the idea of information and psychological influences, and graphs reflecting the distribution in time of the number of persons who adopted the idea of information and psychological influences under certain initial conditions are given. Simulation experiments related to variations of the model parameters in which the dependence of increasing rate of

person's number who adopted the idea of information and psychological effects on the frequency of person's contacts sharing these ideas is studied.

Keywords: system-dynamic modeling, imitating simulation, simulation experiments, information and psychological influence, mass consciousness.

For citation. MINAEV, Vladimir A.; VAITC, Ekaterina V.; GRACHEVA, Yulia V. Simulation Experiments with the Model of Information-Psychological Influences on Mass Consciousness. IT Security (Russia), [S.l.], v. 24, n. 2, p. 61-71, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/106>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.07.

Введение

Выявление особенностей информационно-психологических воздействий на массовое сознание, их прогнозирование и управление ими является одним из важнейших аспектов обеспечения информационно-психологической безопасности общества и его граждан.

Определим информационно-психологическую безопасность как состояние защищенности отдельных лиц и (или) групп лиц от негативных информационно-психологических воздействий (ИПВ) и связанных с ним жизненно важных интересов личности, общества и государства в информационной сфере [1].

Для обеспечения информационно-психологической безопасности необходимо четко понимать структуру и цели информационно-психологических воздействий, особенности их динамики, выделять множество факторов, критично влияющих на данные процессы.

На сегодняшний день накоплен определенный опыт исследований в рамках данной проблематики [2-5], включающий современные математические модели управления массовым сознанием. Он позволяет перейти к имитационному моделированию, предоставляющему широкий спектр возможностей для практического изучения ИПВ.

Авторами для исследования влияния ИПВ на массовое сознание выбран метод системно-динамического моделирования, который позволяет воспроизводить динамику ИПВ в зависимости от состава и структуры элементов моделируемой системы. Системно-динамическое моделирование позволяет делать углубленные и детальные оценки и прогнозы влияния ИПВ на массовое сознание в зависимости от динамики различных факторов. Это направление обосновал и реализовал на моделях сложных социально-экономических явлений Дж. Форрестер в 1950-х годах [6]. Основными переменными в указанных моделях выступают уровни и темпы. Первые отражают возникающее накопление внутри системы, вторые – динамику уровней за один интервал времени. Для описания системно-динамических моделей применяются системные потоковые диаграммы [6].

В части применения метода системно-динамического моделирования для исследования систем обеспечения информационной безопасности накоплен определенный опыт у разных зарубежных научных коллективов [7-17]. Исследование динамики ИПВ на массовое сознание, приведенное в статье, опирается на полученные зарубежными научными коллективами подходы и особенности применения метода системно-динамического моделированию к решению проблем информационной безопасности.

Описание системно-динамической модели

Опишем графически системно-динамическую модель влияния ИПВ на массовое сознание (рис. 1), исходя из результатов, приведенных в работах [2-4]. Условные обозначения, используемые в модели, приведены в таблице 1.

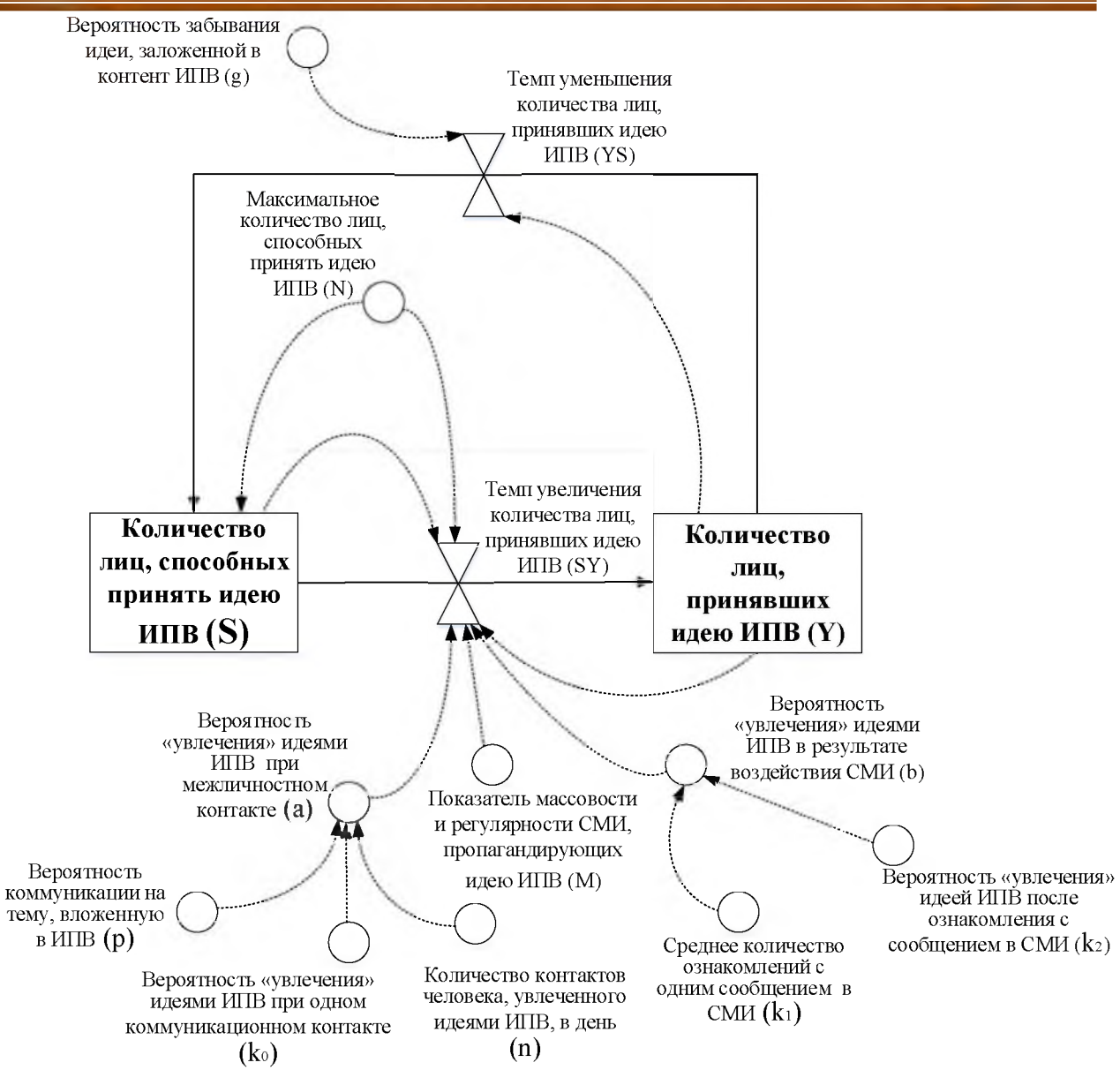


Рис. 1. Системно-динамическая модель влияния ИПВ на массовое сознание
Fig. 1. System dynamics model of the impact of IPV on mass consciousness

Таблица 1 – Условные обозначения, используемые в модели

№ п/п	Условное обозначение элемента	Название элемента
1.	*Y*	Количество лиц, принявших идею ИПВ
2.	*N*	Максимально возможное количество лиц, способных принять идею ИПВ
3.	*S*	Количество лиц, способных принять идею ИПВ
4.	*YS*	Темп уменьшения количества лиц, принявших идею ИПВ
5.	*SY*	Темп увеличения количества лиц, принявших идею ИПВ
6.	*a*	Вероятность «увлечения» идеями, заложенными в контент ИПВ, одного человека за единичный интервал времени при межличностном контакте
7.	*b*	Вероятность «увлечения» идеями, заложенными в контент ИПВ, одного человека за единичный интервал времени, в результате

		воздействия СМИ
8.	M	Показатель массовости и регулярности СМИ, пропагандирующих идею ИПВ
9.	g	Вероятность забывания идеи, заложенной в контент ИПВ, за единичный интервал времени
10.	p	Вероятность коммуникации на тему, вложенную в ИПВ
11.	k_0	Вероятность «увлечения» идеями ИПВ при одном коммуникационном контакте
12.	k_1	Среднее количество ознакомлений с одним сообщением в СМИ за единичный интервал времени
13.	k_2	Вероятность «увлечения» новой идеей из контента ИПВ после ознакомления с сообщением в СМИ
14.	n	Количество контактов человека, увлеченного идеями ИПВ, в день

Построенная системно-динамическая модель описывается следующей системой уравнений:

$$\begin{cases} \dfrac{dS}{dt} = YS(\mathrm{t}) - SY(\mathrm{t}) \\ \dfrac{dY}{dt} = SY(\mathrm{t}) - YS(\mathrm{t}) \\ SY(\mathrm{t}) = \dfrac{S \cdot (a \cdot \mathrm{Y} + M \cdot \mathrm{b})}{N} \\ YS(\mathrm{t}) = g \cdot \mathrm{Y} \\ \mathrm{a} = p \cdot k_0 \cdot n \\ \mathrm{b} = k_1 \cdot k_2 \end{cases}$$

Первое уравнение системы определяет скорость изменения количества лиц, способных принять идею ИПВ. Второе уравнение системы определяет скорость изменения количества лиц, принявших идею ИПВ. Третье уравнение системы определяет темп увеличения количества лиц, принявших идею ИПВ. Четвертое уравнение системы определяет темп уменьшения количества лиц, принявших идею ИПВ. Пятое и шестое уравнения системы определяют значения вероятностей «увлечения» идеями, заложенными в контент ИПВ при межличностном контакте и в результате воздействия СМИ, соответственно.

Данная системно-динамическая модель позволяет проигрывать различные сценарии развития процессов ИПВ и исследовать влияние на них параметров модели.

Вот только некоторые вопросы, на которые может ответить модель:

- Насколько увеличится (уменьшится) количество лиц, принявших идею ИПВ при увеличении (уменьшении) вероятности «увлечения» идеями ИПВ при межличностном контакте?
- Насколько увеличится (уменьшится) количество лиц, принявших идею ИПВ при увеличении (уменьшении) вероятности «увлечения» идеями ИПВ в результате воздействия СМИ?
- Каково должно быть количество контактов в день лица, увлеченного идеями ИПВ, для достижения количеством тех, кто принял идею ИПВ, до определенного значения за заданный интервал времени?
- Как влияет вероятность забывания идеи, заложенной в контент ИПВ, на количество лиц, принявших идею ИПВ?

- Через какое время, при определенных параметрах модели, количество лиц, принявших идею ИПВ, достигнет критического значения?

В качестве платформы для программной реализации построенной модели выбрано перспективное решение в области имитационного моделирования AnyLogic [18]. Общий вид интерфейса модели приведен на рис. 2.

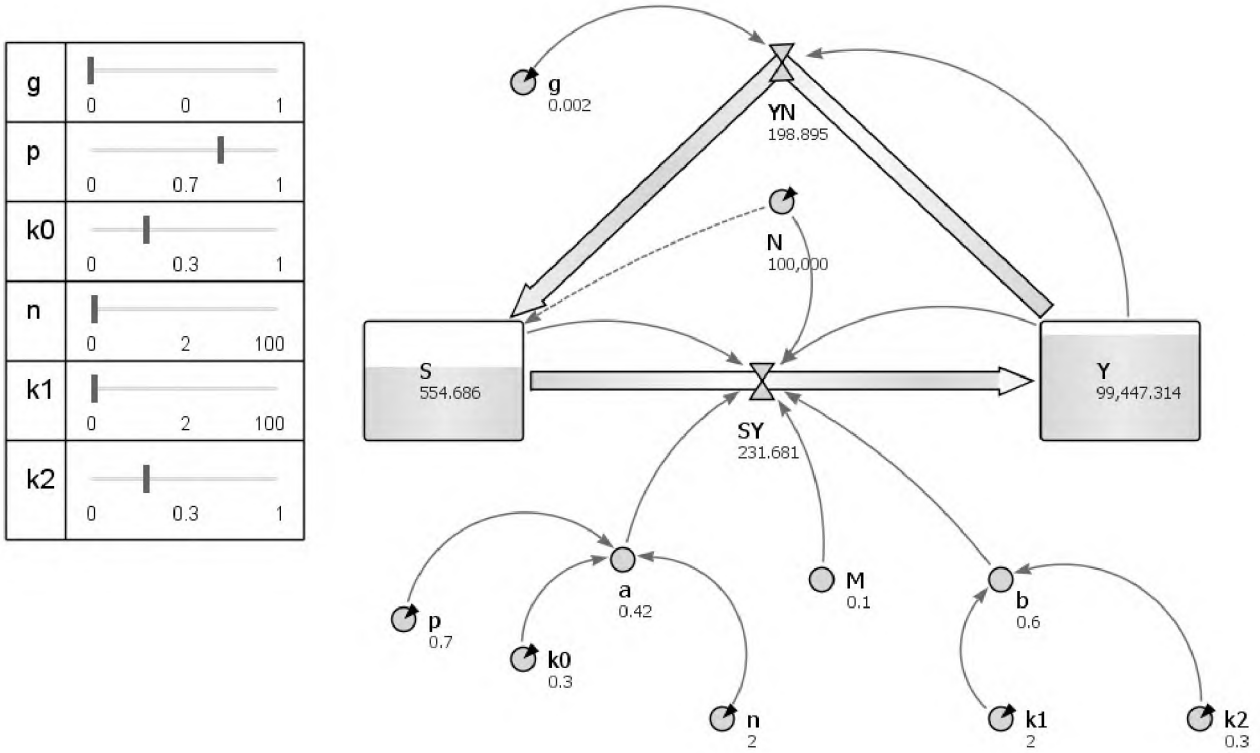


Рис. 2. Общий вид интерфейса модели влияния ИПВ на массовое сознание
Fig. 2. The interface general view of the IPV effects model on the consciousness of mass

Панель управления, приведенная на рис. 2, представляет собой совокупность “бегунков”, каждый из которых закреплен за соответствующим параметром модели. В процессе моделирования можно изменять их значения с помощью “бегунков”, наглядно воспроизводя и отображая происходящие процессы ИПВ.

Приведем, для примера, временные графики, отражающие динамику количества лиц, способных принять идею ИПВ (S), и принявших идею ИПВ (Y), (рис. 3) при следующих начальных условиях:

$$\{\mathrm{p}=0,7; k_0=0,3; n=2; \mathrm{k}_1=2; \mathrm{k}_2=0,3; \mathrm{M}=0,1; \mathrm{g}=0,002; S(0)=100000, \mathrm{Y}(0)=2\}$$

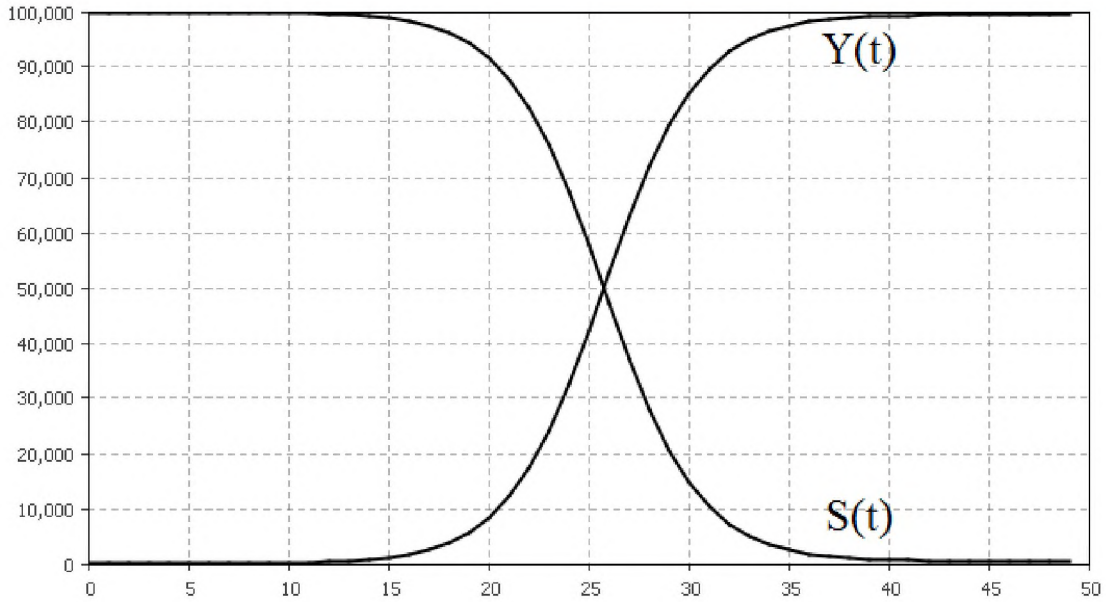


Рис. 3. Динамика количества лиц, способных принять идею ИПВ (S) и принявших её (Y)
Fig. 3. Dynamics of the number of persons able to accept the idea IPV (S) and accepted it (Y)

Рис. 3 демонстрируют, что динамика лиц, принявших идею ИПВ, представляется в виде S-образной кривой, концепцию которой Э. Роджерс [19] описывает следующим образом: «Сначала всего несколько индивидов принимают новую идею, затем инновация принимается большим количеством индивидов, и, наконец, темпы принятия замедляются». Таким образом, первый этап распространения идеи ИПВ представляет собой медленное нарастание до определенного порогового уровня, далее, на втором этапе, наблюдается взрывная фаза распространения ИПВ, и, наконец, на третьем этапе наступает насыщение, при котором лица, принявшие идею ИПВ, контактируют преимущественно друг с другом и процесс информационного «заражения» существенно замедляется.
Рис. 4 демонстрирует распределение (в %) во времени количества лиц, принявших идею ИПВ. Можно выделить пять их категорий: новаторы, ранние принимающие, раннее большинство, позднее большинство, отстающие.

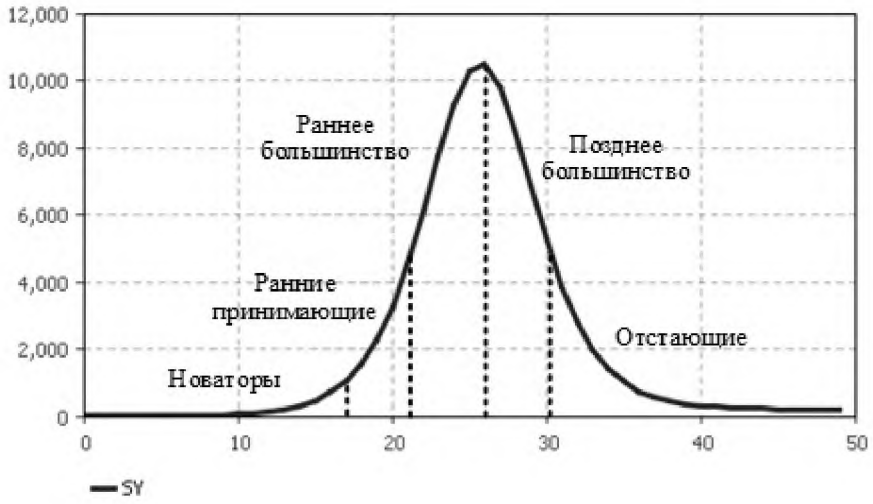


Рис. 4. Распределение (в %) во времени количества лиц, принявших идею ИПВ
Fig. 4. Distribution (%) time number of persons have accepted the idea IPV

Эксперимент по варьированию количества контактов лица, увлеченного идеями ИПВ

Приведем результаты имитационного эксперимента (таблица 2), в котором исследуется зависимость темпа увеличения количества новых членов общества, принявших идею ИПВ, от частоты контактов лиц, уже разделяющих идеи ИПВ.
Иными словами, изучим влияние количества контактов лица, увлеченного идеями ИПВ, в день (n) на темп принятия идеи ИПВ (SY) новыми индивидами.

Таблица 2 - Матрица планирования эксперимента

Номер эксперимента	**Значение варьируемого параметра, *n***	**Темп принятия идеи, *SY***
1.	2	$SY_1(t)$
2.	3	$SY_2(t)$
3.	4	$SY_3(t)$
4.	5	$SY_4(t)$
5.	6	$SY_5(t)$
6.	7	$SY_6(t)$
7.	8	$SY_7(t)$

Начальные значения других параметров модели определены следующим образом:

$$\begin{cases} \mathrm{p} = 0,4 \\ k_0 = 0,3 \\ \mathrm{k}_1 = 2 \\ \mathrm{k}_2 = 0,3 \\ \mathrm{M} = 0,1 \\ \mathrm{g} = 0,002 \\ S(0) = 100000 \\ \mathrm{Y}(0) = 1 \end{cases}$$

На рис. 5 приведены временные графики, полученные по результатам эксперимента
Выводы по результатам эксперимента: проведенный эксперимент выявил существенную зависимость между темпом принятия идеи ИПВ новыми членами сообщества и частотой контактов лиц, уже увлеченных идеями ИПВ, в день (n). Так, при двух контактах в день ($n=2$), темп принятия идеи ИПВ достигнет своего максимума через 48 дней, а уже, например, при шести контактах в день ($n=6$) темп принятия идеи ИПВ достигнет своего максимума в три раза быстрее – уже через 16 дней.

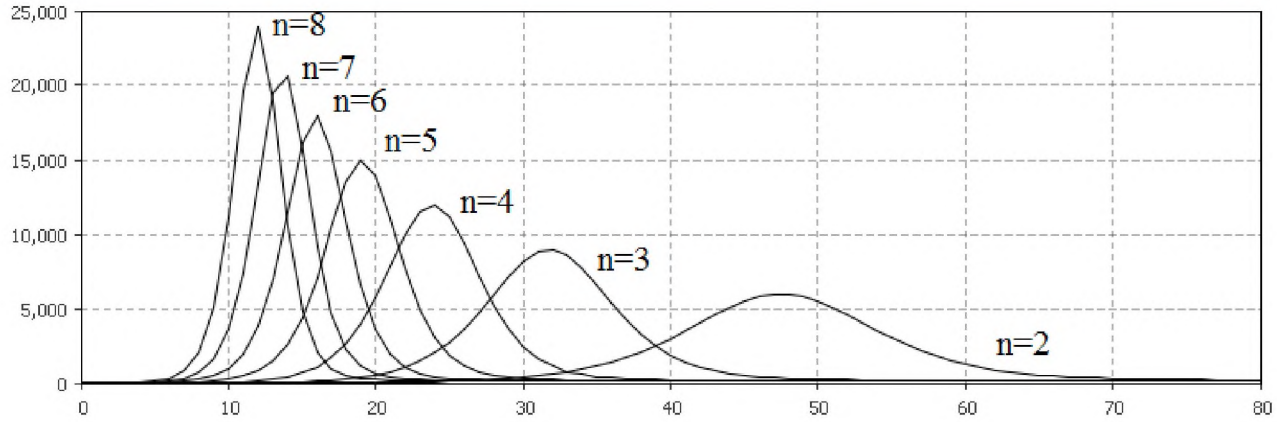


Рис. 6. Темп принятия идеи ИПВ в зависимости от частоты контактов лиц, увлеченных идеями ИПВ, в день

Fig. 6. The rate of adoption of the ideas of IPV depending on the frequency of contacts of persons fascinated by the ideas of IPV in the day

Заключение

1. На сегодняшний день накоплен достаточный опыт в области исследования ИПВ на массовое сознание, опираясь на который можно строить весьма сложные имитационные модели, предоставляющие большие возможности для изучения различных процессов, связанных с ИПВ. Метод системной динамики, выбранный для реализации такого типа моделей, позволяет делать оперативные, реалистичные оценки и прогнозы влияния ИПВ на массовое сознание в зависимости от различных факторов.

2. Реализация построенной системно-динамической модели влияния ИПВ на массовое сознание в программной среде Anylogic позволяет наглядно отслеживать влияние различных важных параметров на происходящие процессы “заражения” общества и оценивать его в реальном масштабе времени.

Дальнейшим развитием видится детализация системно-динамической модели, заключающаяся в добавлении в нее новых уровней и темпов, уточнении комплекса факторов и проведении новой серии имитационных экспериментов с различными комбинациями факторов.

СПИСОК ЛИТЕРАТУРЫ:

1. Ковалева Н.Н. Информационное право России. Учебное пособие. М., 2007. – 360с.
2. Минаев В.А., Дворянкин С.В. Моделирование динамики информационно-психологических воздействий на массовое сознание. Вопросы кибербезопасности, 2016. № 5. – С. 56-64.
3. Минаев В.А. Исследование модели динамики деструктивных информационно-психологических воздействий на массовое сознание. Безопасность информационных технологий, 2016. № 4. – С. 52-58.
4. Минаев В.А. Овчинский А.С., Скрыль С.В., Тростянский С.Н. Как управлять массовым сознанием: современные модели. 2013. - 200 с.
5. Минаев. В.А. Моделирование процессов развития устойчивого туризма. Научный журнал «Сервис в России и за рубежом». Т.8, № 9 (56), 2014. – С 140-149.
6. Форрестер Д. Основы кибернетики предприятия (индустриальная динамика). Москва: Прогресс, 1971. – 340с.
7. Liu Wei, Cui Yong-Feng, Li Ya. Information Systems Security Assessment Based on System Dynamics. International Journal of Security and Its Applications. 2015. Vol. 9. No. 2. Pp. 73-84.

8. Ae Chan Kim, Su Mi Lee, Dong Hoon Lee. Compliance Risk Assessment Measures of Financial Information Security Using System Dynamics. International Journal of Security and Its Applications. 2012. Vol. 6. No. 4. Pp. 191-200.
9. Trcek D. Using System Dynamics for Managing Risks in Information Systems. Wseas transactions on information science & Applications. 2008. Vol. 2. No. 5. Pp. 175-180.
10. Jose M. Sarriegi, Javier Santos, Jose M. Torres, David Imizcoz, Angel L. Plandolit. Modeling Security Management of Information Systems: Analysis of a Ongoing Practical Case. 24th International Conference of the System Dynamics Society. Nijmegen, The Netherlands. 2016. Vol. 1. Pp. 3196-3209.
11. Derek L. Nazaretha, Jae Choib. A system dynamics model for information security management. Information & Management. 2015. Vol. 52. No. 1. Pp. 123-134.
12. Cappelli D.M., Desai A.G., Moore A.P., Shimeall T.J., Weaver E.A., Bradford B.J. Management and Education of the Risk of Insider Threat (MERIT): System Dynamics Modeling of Computer System Sabotage, Carnegie-Mellon University. Software Engineering In Institute. CERT Program. , Pittsburg, 2006. – 34p.
13. Mundie D.A., Collins M.L., Moore A.P. A System Dynamics Model for Investigating Early Detection of Insider Threat Ris, Carnegie-Mellon University. Software Engineering Institute. CERT Program, Pittsburg, 2013. – 24p.
14. Yang S.C., Yi-Lu Wang. System Dynamics Based Insider Threats Modeling. International Journal of Network Security & Its Applications (IJNSA). 2011. Vol. 3. No. 3. Pp 1-14.
15. Application and Management of Cybersecurity Measures for Protection and Control Systems, CIGRE, Joint Working Group B5/D2.46, 2014. – 121p.
16. Behara R.S., C. Derrick Huang, Qing Hu. A System Dynamics Model of Information Security Investments. Journal of Information System Security. 2010. Vol. 6. No. 2. Pp. 1572-1583.
17. Julia H. Allen , Noopur Davis. Measuring Operational Resilience Using the CERT Resilience Management Model, Carnegie Mellon Universit, 2010. – 71p.
18. Каталевский Д.Ю. Основы имитационного моделирования и системного анализа в управлении: учебное пособие; 2-е изд., перераб. и доп. — М.: Издательский дом «Дело» РАНХиГС, 2015. — 496 с.
19. Rogers E. Diffusion of Innovations. – 4 ed. – New York: Free Press, 1995. – Pp. 204-251.

REFERENCES:

[1] Kovaleva N.N. Information law of Russia. Tutorial. Moscow, 2007. – 360p. (in Russia).
[2] Minaev V.A., Dvoryankin S.V. Modeling the dynamics of information and psychological influence on mass consciousness. Questions of cybersecurity, 2016. № 5. Pp. 56-64. (in Russia).
[3] Minaev V.A. Investigation of the model of dynamics of destructive information and psychological influences on mass consciousness. IT Security, 2016. № 4. Pp. 52-58. (in Russia).
[4] Minaev V.A. Ovchinsky A.S., Skryl S.V., Trostyanskiy S.N. How to control the mass consciousness: modern models. 2013. – 200p. (in Russia).
[5] Minaev. V.A. Modeling the processes of sustainable tourism development. Scientific magazine "Service in Russia and abroad." T.8, No. 9 (56), 2014. Pp. 140-149. (in Russia).
[6] Forrester D. Fundamentals of enterprise cybernetics (industrial dynamics). Moscow: Progress, 1971. – 340p. (in Russia).
[7] Liu Wei, Cui Yong-Feng, Li Ya. Information Systems Security Assessment Based on System Dynamics. International Journal of Security and Its Applications. 2015. Vol. 9. No. 2. Pp. 73-84.
[8] Ae Chan Kim, Su Mi Lee, Dong Hoon Lee. Compliance Risk Assessment Measures of Financial Information Security Using System Dynamics. International Journal of Security and Its Applications. 2012. Vol. 6. No. 4. Pp. 191-200.
[9] Trcek D. Using System Dynamics for Managing Risks in Information Systems. Wseas transactions on information science & Applications. 2008. Vol. 2. No. 5. Pp. 175-180.

[10] Jose M. Sarriegi, Javier Santos, Jose M. Torres, David Imizcoz, Angel L. Plandolit. Modeling Security Management of Information Systems: Analysis of a Ongoing Practical Case. 24th International Conference of the System Dynamics Society. Nijmegen, The Netherlands. 2016. Vol. 1. Pp. 3196-3209.
[11] Derek L. Nazaretha, Jae Choib. A system dynamics model for information security management. Information & Management. 2015. Vol. 52. No. 1. Pp. 123-134.
[12] Cappelli D.M., Desai A.G., Moore A.P., Shimeall T.J., Weaver E.A., Bradford B.J.. Management and Education of the Risk of Insider Threat (MERIT): System Dynamics Modeling of Computer System Sabotage, Carnegie-Mellon University. Software Engineering In Institute. CERT Program. , Pittsburg, 2006. – 34p.
[13] Mundie D.A., Collins M.L., Moore A.P. A System Dynamics Model for Investigating Early Detection of Insider Threat Ris, Carnegie-Mellon University. Software Engineering Institute. CERT Program, Pittsburg, 2013. – 24p.
[14] Yang S.C., Yi-Lu Wang. System Dynamics Based Insider Threats Modeling. International Journal of Network Security & Its Applications (IJNSA). 2011. Vol. 3. No. 3. Pp 1-14.
[15] Application and Management of Cybersecurity Measures for Protection and Control Systems, CIGRE, Joint Working Group B5/D2.46, 2014. – 121p.
[16] Behara R.S., C. Derrick Huang, Qing Hu. A System Dynamics Model of Information Security Investments. Journal of Information System Security. 2010. Vol. 6. No. 2. Pp. 1572-1583.
[17] Julia H. Allen , Noopur Davis. Measuring Operational Resilience Using the CERT Resilience Management Model, Carnegie Mellon University, 2010. – 71p.
[18] Katalevsky D.Yu. Fundamentals of simulation and system analysis in management: a textbook; 2 nd ed., Revised. and additional. - M .: Publishing house "Delo" RANHiGS, 2015. - 496p. (in Russia).
[19] Rogers E. Diffusion of Innovations. - 4 ed. - New York: Free Press, 1995. Pp. 204-251.

Поступила в редакцию – 22 февраля июля 2017 г. Окончательный вариант – 23 мая 2017 г.
Received – February 22, 2017. The final version – May 23, 2017.

Иван В. Нечта
Сибирский государственный университет телекоммуникаций и информатики,
ул. Кирова, 86, Новосибирск, 630102, Россия,
e-mail: ivannechta@gmail.com, ORCID iD 0000-0003-0361-2742

АНАЛИЗ УСТОЙЧИВОСТИ ДИНАМИЧЕСКИХ ВОДЯНЫХ ЗНАКОВ

DOI: http://dx.doi.org/10.26583/bit.2017.2.08

Аннотация. В данной работе рассмотрена ранее известная схема внедрения динамических водяных знаков Radix-n, применяемая для борьбы с нелегальным использованием программного обеспечения. Согласно схеме водяной знак представляет собой динамическую структуру данных (граф), которую программа создает в памяти компьютера в процессе выполнения. Топология такого графа определяет скрытую информацию, например, сведения об авторе. Эта информация может быть извлечена для установления авторства в рамках судебных разбирательств. В ходе данного исследования было установлено, что рассматриваемая схема, считавшейся ранее одной из наиболее надежных, обладает рядом особенностей, позволяющих злоумышленнику выявить этап построения водяного знака в программе и, соответственно, его исказить или удалить. Автором данной статьи показан недостаток схемы Radix-n, который заключается в том, что с помощью сведений, полученных от перехватчика Api-функций выделения динамических блоков памяти, можно выявить динамические структуры данных программы, среди которых присутствует водяной знак. Анализ содержимого оперативной памяти позволяет обнаружить указатели на динамически созданные объекты (массивы, переменные, экземпляры классов и т.д.) программы. Различные динамические объекты в памяти, связанные между собой указателями образуют структуры данных программы (списки, стеки, деревья и иные графы), в том числе и водяной знак. В ходе проведенных экспериментов установлено, что в подавляющем числе случаев количество структур данных в программах невелико, что повышает вероятность реализации успешной атаки злоумышленником. В настоящей работе также предложен алгоритм поиска компонент связности графа с линейной трудоемкостью, в случае, когда количество вершин порядка 106. На основе полученных в эксперименте результатов была предложена новая схема внедрения водяных знаков, являющаяся устойчивой к разработанной атаке. В предлагаемой схеме используется граф для хранения водяного знака, где связи между узлами реализованы с помощью уникальных сигнатур. Предложено использовать шифрование содержимого узлов графа (кроме сигнатур) для того, чтобы злоумышленник не смог выявить наличие связи между узлами.

Ключевые слова: динамические водяные знаки, стеганография исполняемых файлов.

Для цитирования. НЕЧТА, Иван В. Анализ устойчивости динамических водяных знаков. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 72-81, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/107>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.08.

Ivan V. Nechta
Sibirian state university of telecommunications and informatic sciences,
Kirova st., 86, Novosibirsk, 630102, Russia,
e-mail: ivannechta@gmail.com, ORCID iD 0000-0003-0361-*2742*

Robustness analysis of dynamic watermarks

DOI: http://dx.doi.org/10.26583/bit.2017.2.08

Abstract. In this paper we consider previously known scheme of dynamic watermarks embedding (Radix-n) that is used for preventing illegal use of software. According to the scheme a watermark is dynamic linked data structure (graph), which is created in memory during

program execution. Hidden data, such as information about author, can be represented in a different type of graph structure. This data can be extracted and demonstrated in judicial proceedings. This paper declared that the above mentioned scheme was previously one of the most reliable, has a number of features that allows an attacker to detect a stage of watermark construction in the program, and therefore it can be corrupted or deleted. The author of this article shows the weakness of Radix-N scheme, which consists in the fact that we can reveal dynamic data structures of a program by using information received from some API-functions hooker which catches function calls of dynamic memory allocation. One of these data structures is the watermark. Pointers on dynamically created objects (arrays, variables, class items, etc.) of a program can be detected by content analysis of computer`s RAM. Different dynamic objects in memory interconnected by pointers form dynamic data structures of a program such as lists, stacks, trees and other graphs (including the watermark). Our experiment shows that in the vast majority of cases the amount of data structure in programs is small, which increases probability of a successful attack. Also we present an algorithm for finding connected components of a graph with linear time-consuming in cases where the number of nodes is about 106. On the basis of the experimental findings the new watermarking scheme has been presented, which is resistant to the proposed attack. It is offered to use different graph structure representation of a watermark, where edges are implemented using unique signatures. Our scheme uses content encrypting of graph nodes (except signature), so an attacker is not able to reveal links between nodes and find the watermark.

Keywords: dynamic watermarking, steganography of executables.

For citation. NECHTA, Ivan V. Robustness Analysis of Dynamic Watermarks. IT Security (Russia), [S.l.], v. 24, n. 2, p. 72-81, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/107>. Date accessed: 23 june. 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.08.

Введение

Современные производители программного обеспечения большое внимание уделяют методам защиты программ от нелегального использования. Под «нелегальным» понимается использование программы без покупки лицензионного ключа. Существуют множество методов защиты программы от дизассемблирования или редактирования, например [1-3], однако злоумышленнику зачастую удается найти место в программе, где производится проверка правильности лицензионного ключа и либо исказить алгоритм проверки, либо создать простейший генератор ключей (т.н. keygen). В результате успешной атаки злоумышленник выкладывает в общий доступ программу, лишенную всякой защиты от нелегального использования.

Еще одной угрозой, с которой сталкивается производитель, является незаконное использование программы или ее компонентов недобросовестными конкурентами в составе чужих программ (аналог плагиата). Зачастую автору бывает трудно доказать некорректное заимствование компонентов своих программ в чужом программном продукте.

Авторы статей [4-5] предлагают, чтобы для защиты от подобных проблем производитель встраивал некую секретную информацию (цифровой водяной знак, далее ЦВЗ) в свой продукт, которая содержит сведения об «истинном» авторе. В случае обнаружения некорректного заимствования эта информация может быть извлечена и использована в ходе судебных разбирательств (см. работу [6]).

Очевидно, что водяной знак должен обладать некоторой устойчивостью[1], чтобы злоумышленник не смог его исказить или удалить из программы. В ряде случаев производители явно показывают свое авторство, например, в виде строк об авторстве в меню «О программе», в таком случае злоумышленнику легко найти и исказить ЦВЗ. В

[1] Здесь и далее под «устойчивостью» понимается свойство, обеспечивающее невозможность обнаружения или искажения ЦВЗ в программе.

других случаях информация может быть встроена в особом виде, скрытом от стороннего наблюдателя, когда злоумышленник не может гарантировать ни наличия, ни отсутствия ЦВЗ. Рассмотрим более подробно методы внедрения скрытых ЦВЗ.

Существующие ЦВЗ можно условно разделить на две основные категории:

- статические ЦВЗ, описанные в работах [7-9], представляют данные в виде констант или инструкций программы, фрагментов неиспользуемого кода (т.н. «мертвый код»);
- динамические ЦВЗ, предложенные в работе [10], представляют данные в виде динамической структуры данных, создаваемой в оперативной памяти компьютера на этапе выполнения программы.

Анализ статей [11-13] показал, что наиболее устойчивыми к обнаружению являются динамические ЦВЗ, т.к. анализ работы программы производится пошаговым выполнением команд (т.н. трассировкой), что занимает достаточно большое время. В таких условиях считается затруднительным выявление динамической структуры данных в памяти, отвечающей за водяной знак.

В работе [10] предлагается схема Radix-n динамического ЦВЗ, представленная на рис. 1. Структура водяного знака представляет собой граф. Имеется ключевая вершина, по которой этот граф может быть найден в памяти (например, содержит некоторую сигнатуру, известную только автору). Топология графа определяет содержимое ЦВЗ.

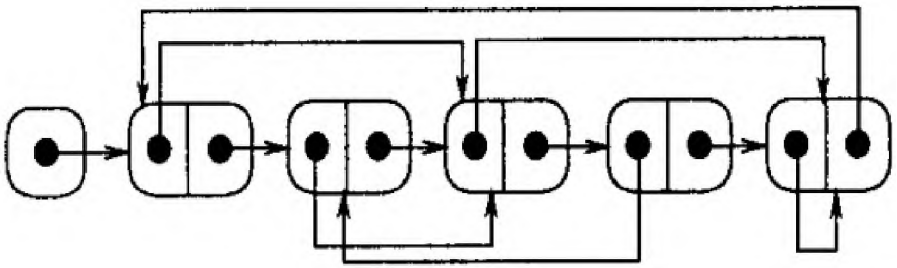


Рис. 1. Цифровой водяной знак по схеме Radix-n
Fig. 1. A digital watermark according to the scheme Radix-n

В той же работе рассматривались возможные атаки, но они требовали наличие исходных кодов программ.

Строго говоря, существует два типа атак:

- статические, которые имеют дело с анализом исходных кодов программ или, что чаще встречается, с анализом через дизассемблирование программы;
- динамические, которые предполагают анализ работы программы в ходе ее выполнения.

Считается, что статические атаки являются менее эффективными, чем динамические. Чаще всего злоумышленник имеет дело с готовым исполняемым файлом (исходные коды ему не доступны[2]) и статический анализ через дизассемблирование не дает полной картины происходящих процессов в программе. Динамический анализ, с одной стороны, дает качественную картину действий программы (Api вызовы, доступ к жесткому диску или сети), и с другой стороны, позволяет отсеять большой объем неиспользуемого кода в самом начале анализа.

По утверждениям авторов статьи [10], в процессе работы программы найти место построения ЦВЗ не представляется возможным, т.к. для этого требуется проанализировать содержимое всей динамической памяти. Во-первых, не ясно как отличить указатели от

[2] Речь идет только о файлах формата portable executable (exe, dll, …). Программы, выполняемые на Java Virtual Machine, с легкостью дизассемблируются и восстанавливаются до высокоуровневого исходного кода.

иных данных, а во-вторых, программа может создавать множество аналогичных структур данных, неотличимых от водяного знака, что усложняет задачу анализа.

В настоящей статье предлагается эффективный метод поиска ЦВЗ в исполняемых файлах. Показано, что схема Radix-n имеет ряд недостатков, которые способствуют в подавляющем большинстве случаев обнаружению времени и места построения ЦВЗ в программе.

Описание предлагаемого анализа.

Основная задача, которая ставится в рамках реализации атаки на ЦВЗ, состоит в нахождении этапа в работе программы, во время которого производится построение водяного знака. Выявив указанный этап, злоумышленник с легкостью определит участок кода, ответственный за построение ЦВЗ и произведет его искажение. Наличие дополнительных механизмов контроля целостности (контроль отсутствия изменений) программы, так и противодействие им, не является предметом данного исследования и должно рассматриваться отдельно.

Рассмотрим следующий пример. Допустим, имеется программа, которая при запуске создает ЦВЗ указанным в работе [10] способом (Radix-n), и затем, продолжает работу. Для обнаружения указанного этапа работы программы злоумышленнику потребуется найти в памяти все графоподобные структуры данных, причем узлы графа представляют собой динамически выделенные области памяти, а дуги представлены в виде указателей.

Будем утверждать, что если количество таких структур, создаваемых программой, менее определенного числа $Count_{gr}$, то злоумышленник может без особых усилий выявить среди них ЦВЗ, при наличии полных сведений о структурах. Предлагаемый метод анализа эффективно может работать при $Count_{gr} = 20$, что связано со сложностью редактирования и анализа программ, позволяющее выполнить искажение ЦВЗ. В ситуации, количество указанных структур более 20, предлагаемая в данной работе атака не реализуема. Также будем называть ЦВЗ программы *графом*, т.к. он представлен в виде ориентированного графа и расположен в динамической памяти процесса.

Теперь определим способы получения сведений о создаваемых программой структур данных в памяти. Известно, что для выделения динамической памяти программисты используют специальные Api-функции. В данной работе рассматривается функция *malloc* библиотеки языка C/C++. Стоит отметить, что иные функции (*calloc*, *new* и т.д.) в свою очередь вызывают *malloc*. Будем использовать перехватчик (см. рис. 2) вызовов Api-функций для получения следующей информации:

- время вызова, определяющее этап работы программы;
- размер выделенной памяти;
- адрес выделенной памяти (необходим для последующего анализа содержимого);
- имя модуля запросившего память (это позволит отсеивать ненужные вызовы системных библиотек Winodws, не относящихся к ЦВЗ).

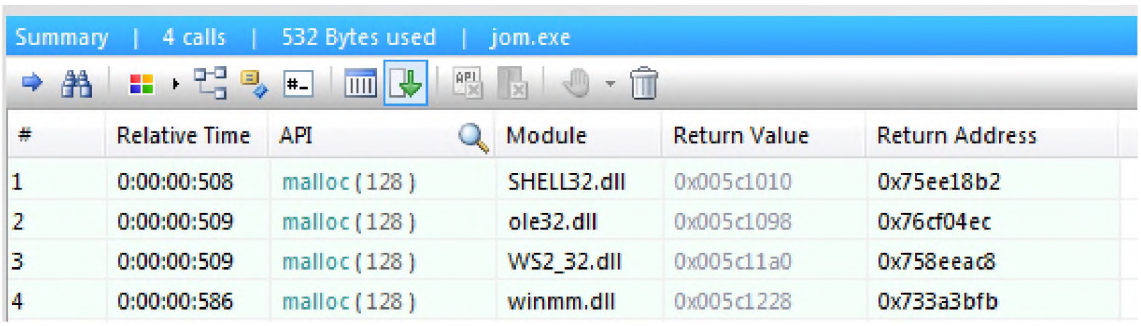


Рис. 2. Перехват вызовов Api-функций программы
Fig. 2. Intercepts calls to the Api functions of the program

Получив все необходимые сведения об узлах графа, рассмотрим способ поиска дуг (указателей). Согласно схеме (рис. 1.) ЦВЗ представляет собой отдельный граф, имеющий одну компоненту связности. Указатели в памяти могут быть найдены следующим образом. Пусть $M = \{m_1, m_2, \dots, m_N\}$ последовательность байт участка памяти, N – размер этого участка в байтах. Указатель представляет собой четверку последовательных байт $\{m_i, m_{i+1}, m_{i+2}, m_{i+3}\}$, где $i \in [1; N-3]$, причем указывать он должен на одну из выделенных ранее областей памяти. В текущей реализации в качестве указателей брались такие, которые указывали не только на начало динамически выделенных блоков памяти, но и в произвольное место (внутри) данных блоков.

Таким образом, во всех блоках памяти, выделенных динамически, производится поиск указателей и выявляются связанные структуры данных.

Постановка эксперимента.

В рамках эксперимента планируется показать, что в подавляющем большинстве программ количество связанных структур данных в памяти процесса не превосходит указанной границы $Count_{gr}$. Для анализа содержимого памяти процесса использовался Дамп[3] памяти (подробнее см. работу [14]), который в дальнейшем анализировался. Программа запускалась и завершалась через 10 секунд после полной загрузки.

Было установлено, что процесс выделения памяти (по количеству и объему) идет наиболее интенсивно именно в указанном периоде. Соответственно, построение ЦВЗ разумнее выполнять на данном этапе работы программы. В действительности построение может происходить на любом этапе, но как было установлено в экспериментах, количество создаваемых структур в этот период будет много меньше, чем при запуске и это будет значительно облегчать задачу злоумышленнику.

Будем рассматривать выделяемые программой в памяти блоки данных как некий глобальный граф. Одной из задач, решаемых в рамках данного исследования, является поиск компонент связности глобального графа. При хранении глобального графа в виде матрицы поиск компонент связности может осуществляться с помощью ранее известных алгоритмов [15]. В нашем случае, когда количество узлов глобального графа может достигать порядка 10^6, хранение и обработка графа в виде двумерного массива невозможна.

Рассмотрим схему хранения и поиска компонент связности графа, разработанную автором настоящей статьи и используемую в данном эксперименте. На рис. 3 показан ориентированный граф из 6 узлов.

Каждой вершине соответствует свой элемент (показан прямоугольником), содержащий номер группы. Изначально число групп равно числу узлов графа. Просматривая каждый узел слева направо объединим две соответствующих группы в одну, если текущий узел А имеет связь с узлом Б и дуга идет в направлении от А к Б.

Объединенная группа получает номер соответствующий минимальному из номеров объединенных групп. После однократного прохода по каждой вершине, мы получаем узлы, объединенные в группы. Количество объединяющих групп равно числу компонент связности глобального графа. В итоге мы получаем алгоритм поиска компонент связности, имеющий линейное время работы и не требующего большого объема памяти.

Ниже идет таблица полученных результатов. Здесь «искомым графом» называется компонента связности, состоящая из двух или более узлов (среди них может быть ЦВЗ). Ранее мы уже говорили о том, что если искомых графов будет менее чем $Count_{gr} = 20$, то злоумышленник без труда сможет найти и исказить водяной знак.

[3] Дамп – сохраненное в виде файла содержимое оперативной памяти процесса в некоторый момент времени выполнения (в нашем случае на 10 сек. после загрузки).

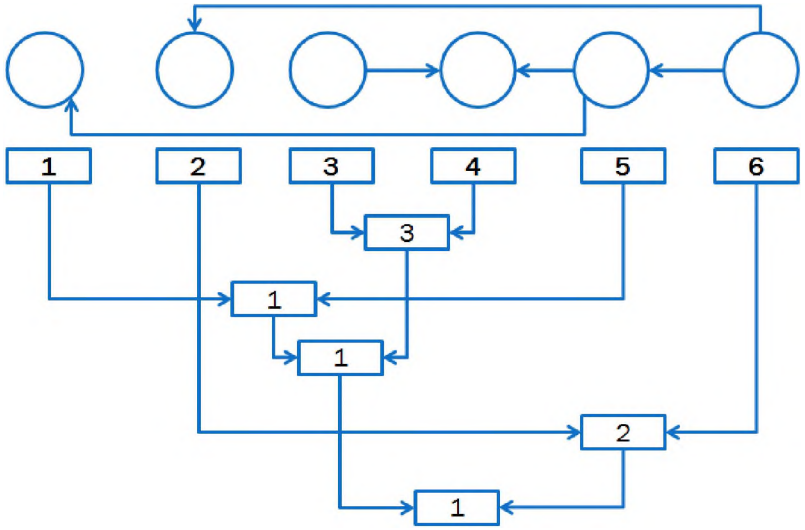


Рис. 3. Схема хранения и поиска компонент связности графа
Fig. 3. The scheme of storage and retrieval of connected components of the graph

Таблица 1. Результаты экспериментальных исследований

Название программы	Количество		
	узлов в глобальном графе	не изолированных узлов глобального графа[4]	искомых графов
WinRar (v.5.00)	159	109	4
Far (v. 3.0.4438)	26	0	0
Gimp (v. 2.8.18)	768667	55412	5030
Chrome (v. 55.0.2883.87)	5222	4116	4
TeXworks (v. 0.4.4)	42737	23523	1655
Word 2007	86	40	2
Excel 2007	52	1	1
Power Point 2007	45	1	1
QtCreator (v. 4.1.0)	466	170	5
Notepad ++ (v. 7.2.2)	22	1	1

Из таблицы видно, что в подавляющем большинстве программ искомых графов менее $Count_{gr}$, это означает, что ЦВЗ может быть обнаружен при помощи предлагаемой атаки. В ходе анализа представленного ряда программ было установлено следующее:

1. При многократных запусках одной и той же программы последовательность вызовов Api-функций оставалась неизменной. Эта особенность дает злоумышленнику возможность отслеживать этап построения ЦВЗ с последующим его искажением.
2. Более 90% блоков памяти (по количеству) выделяется на этапе запуска программы, что делает этот этап работы приложения перспективным для построения ЦВЗ.
3. Содержимое Дампа, за исключением адресов (указателей), практически не меняется от запуска к запуску.
4. Адреса указателей выровнены на границу в 4 байта. Как правило, это связано с тем, что компиляторы по умолчанию осуществляют такое выравнивание (хотя существует способы это выравнивание отключить для оптимизации расходов памяти).

[4] Здесь посчитаны узлы, имеющие хотя бы одну исходящую дугу.

5. При выходе из функции программы, в которой создавались локальные динамические объекты, будет производиться их удаление (для предотвращения так называемой «утечки памяти»). Соответственно, в нашем анализе мы обнаружим, что после запроса на выделение динамической памяти сразу же следует запрос на ее освобождение. В этой связи в Таблице 1 появляется множество дополнительных узлов, которые могут иметь связи. (Реальное количество существующих узлов значительно меньше, например, для Gimp меньше в 7 раз).

Рассмотрим более подробно работу программы Gimp, имеющей наибольшее количество узлов из анализируемого ряда программ. Проведем дополнительное отсеивание узлов, которые, во-первых, уже были освобождены вызовом Api-функции *free*, а во-вторых, имеют размер узла соответствующий атакуемой схеме ЦВЗ (см. рис. 1., размер узла равен 8 байт, т.к. всего два указателя размером по 4 байта).

Таблица 2. Результаты дополнительного отсева узлов

Название программы	Количество		
	узлов в глобальном графе	не изолированных узлов глобального графа	искомых графов
Gimp	2632	16	8

После дополнительного отсеивания мы также видим, что количество искомых графов менее $Count_{gr}$.Таким образом, можно утверждать, что предложенная в работе [10] схема Radix-n для внедрения динамического ЦВЗ не является устойчивой.
Стоит отметить, что в общем случае мы можем продолжать производить отсев узлов, например, по имени модуля (отсекая вызовы от известных системных библиотек от анализа) или иным способом.

Обсуждение полученных результатов.

В предыдущей главе было показано, что применение схемы Radix-n для внедрения ЦВЗ не гарантирует его устойчивости. В нашем эксперименте мы выявляли графы похожие на динамический ЦВЗ. В ходе эксперимента было показано, что при помощи перехвата Api-функции выделения памяти и поиска указателей злоумышленник может существенно сократить время поиска ЦВЗ. Более того, здесь может быть дополнительно произведен отсев найденных графов по топологии графа. Так, часто используемые в программировании графы типа список, стек, дерево и т. д. могут быть исключены из рассмотрения, как не соответствующие схеме Radix-n, что существенно повышает вероятность успешной атаки на водяной знак.

Очевидно, что рассмотренная схема построения ЦВЗ может быть модифицирована таким образом, чтобы в памяти вместо указателя хранилось некоторое смещение. То есть указатель представлен в виде: Указатель = Константа + Смещение. Смещение хранится в памяти, а Константа может рассчитываться программой. Такая схема, на первый взгляд, не позволяет искать связанные структуры, относящиеся к ЦВЗ. Однако при детальном рассмотрении выясняется, что злоумышленник может успешно определить место хранения указателей (ребер графа), относящихся к ЦВЗ следующим образом.

Шаг 1. Запустить программу и получить дамп памяти D_1;

Шаг 2. Повторно запустить программу и получить дамп памяти D_2;

Шаг 3. Получить Дамп разницы D_3 путем побайтового вычитания дампов D_1 и D_2;

Шаг 4. Отсеять указатели на существующие объекты в памяти. Оставшиеся указатели будут относиться к ЦВЗ.

Теперь рассмотрим действия злоумышленника более подробно. В предыдущей главе мы упоминали про то, что содержимое Дампа, за исключением адресов (указателей),

практически не меняется от запуска к запуску. Известно, что все динамические объекты хранятся в специальной области памяти. Адрес указанной области определяется ОС Windows и при каждом новом запуске программы может меняться. Адреса указателей, являющихся суммой адреса начала динамической области памяти и смещения внутри нее меняются от запуска к запуску программы. Соответственно на шаге 3 мы получим дамп D_3, содержащий только указатели, все остальные элементы динамической памяти (т.к. они одинаковые в D_1 и D_2) обнулятся.

Большая часть указателей не относится к ЦВЗ и указывают на реально существующие объекты в динамической памяти. Такие указатели можно успешно выявить и отсеять. Оставшиеся указатели полностью будут относиться к ЦВЗ, что даст злоумышленнику возможность успешно реализовать атаку.

Таким образом, в данной работе было показано, что внедрение динамического ЦЗВ по схеме Radix-n не является устойчивой. Более того модификация связанная с хранением смещений вместо указателей также не повышает устойчивости.

Предлагаемая схема внедрения ЦВЗ.

В предыдущих главах было показано, что исходная схема внедрения ЦВЗ Radix-n позволяет злоумышленнику отслеживать водяной знак в программе. В данной главе будет предложена новая схема внедрения ЦВЗ, которая будет устойчивой к атаке, предложенной в настоящей работе.

Новая схема внедрения предполагает хранение ЦВЗ в виде графа. В отличие от Radix-n узлы графа имеют такой размер, который наиболее часто используется в программе. Рассмотрим рис. 4, на котором показан график зависимости количества выделяемых блоков памяти от его размера. Здесь приводятся данные для программы Gimp.

Из графика мы видим, что в программе чаще всего (10725 раз) выделяются блоки памяти размером 18 байт. Следовательно, размер узлов графа, хранящего ЦВЗ должен быть таким же. Это позволит избежать поиска узлов графа, хранящего ЦВЗ, по заранее известному размеру (в схеме Radix-n размер узла составляет 8 байт).

Вторым отличием предлагаемой схемы внедрения является способ представления данных в ЦВЗ. Radix-n предполагает хранение данных в виде топологии графа, однако в нашем случае ЦВЗ может храниться в виде зашифрованной последовательности байт в каждом узле графа.

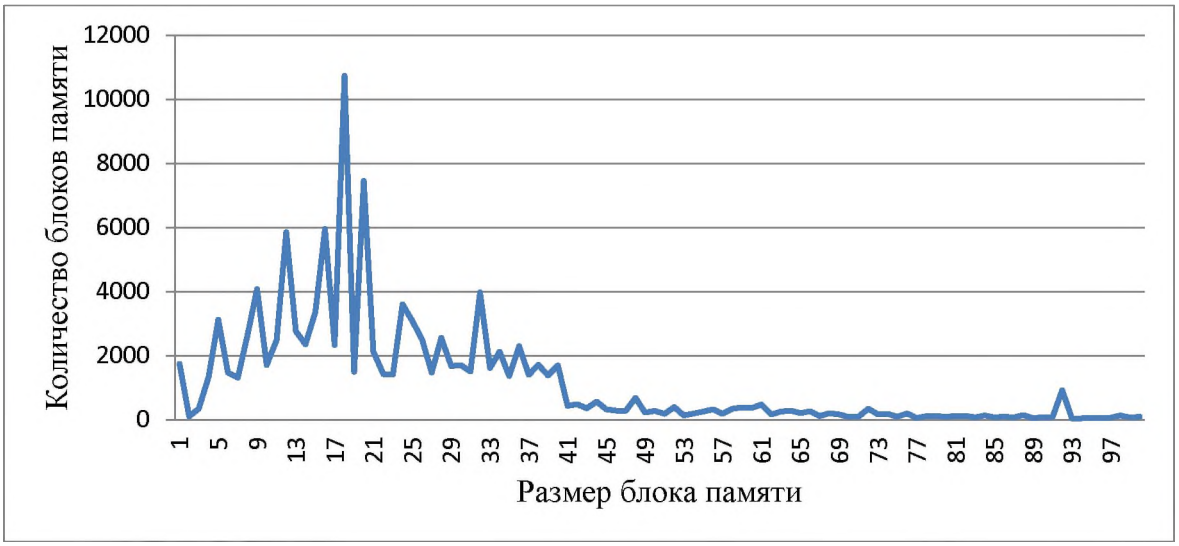


Рис. 4. График зависимости количества блоков памяти от его размера
Fig. 4. The graph of the number of the blocks memory of its size

Третьим отличием нашей схемы является способ связи узлов графа. В связи с тем, что описанная в данной работе атака базируется на поиске связанных указателями узлов графа, предлагается связь реализовывать через специальные сигнатуры. Сигнатурой называется уникальная последовательность байт (фиксированного размера), располагающаяся в начале выделенного блока памяти. Уникальность сигнатуры проверяется на этапе внедрения ЦВЗ. Таким образом, в узле графа, хранящего ЦВЗ, содержится следующая информация:

- сигнатура текущего узла;
- полезные данные ЦВЗ (например, сведения об авторе);
- сигнатура следующего узла (заменяющая собой указатель).

Шифрованию подвергаются полезные данные ЦВЗ и сигнатура следующего узла. Наличие шифрования гарантирует, что злоумышленник не сможет отследить связанность узлов графа. Отказ от использования указателей объясняется тем, что при каждом запуске программы указатели могут менять свое значение, соответственно шифрование должно производиться при каждом запуске программы. Считается недопустимым хранение пароля для шифрования в самой программе. В случае с сигнатурами, шифрование производится единственный раз на этапе внедрения ЦВЗ. Таким образом, пароль в программе хранить не требуется.

В данной работе мы рассмотрели известную схему внедрения Radix-n. Показано, что с помощью метода, предложенного в данной работе, злоумышленник может выявлять ЦВЗ, хранящийся в программе. Также была предложена новая схема внедрения ЦВЗ, которая является устойчивой к данной атаке.

СПИСОК ЛИТЕРАТУРЫ:

1. Алейников С. И., Богатов А. О. Защита программ от дизассемблирования. Труды Института системного программирования РАН. 2006. Т. 11. С. 9-16.
2. Collberg C., Thomborson C. and Low D. A Taxonomy of Obfucating Transformations,Technical Report N148, Department of Computer Science. The University of Auckland, July, 1997.
3. Никольская К. Ю., Хлестов А. Д. Обфускация и методы защиты программных продуктов. Вестник УрФО. Безопасность в информационной сфере. 2015. № 2. С. 7-10.
4. Yarmolik V.N., Portyanko S.S. State of the ART in Software Ownership Protection. Computer Information Systems and Industrial Management Application Editors: Khalid Saeed, Romuald Mosdorf, Olli-Pekka Hilmola. Bialystok, Poland. 2003. pp. 188-195.
5. Collberg C.S., Thomborson C. Watermarking, Tamper-proofing and Obfuscation – Tools for Software protection. Computer Science Technical Report #170. University of Auckland, Aucland, New Zealand. February 10, 2000.
6. Козубенко Ю. В. Защита авторских прав на программы для ЭВМ в уголовном, административном и гражданском судопроизводстве. Wolters Kluwer. Russia, 2009, 344 с.
7. Alexander G. Gounares Fingerprinting Executable Code. Patent US20120317421 A1, date 19.07.2012.
8. Anckaert B. et al. Steganography for executables and code transformation signatures. International Conference on Information Security and Cryptology. Springer Berlin Heidelberg. 2004. pp. 425-439.
9. Zaidan A. A. et al. Novel approach for high (secure and rate) data hidden within triplex space for executable file. Scientific Research and Essays. 2010. vol. 5. no. 15. pp. 1965-1977.
10. Collberg C., Thomborson C. Software watermarking: Models and dynamic embeddings. Proceedings of the 26th ACM SIGPLAN-SIGACT symposium on Principles of programming languages. ACM. 1999. pp. 311-324.
11. Moser A., Kruegel C., Kirda E. Limits of static analysis for malware detection. Computer security applications conference, 2007. ACSAC 2007. Twenty-third annual. IEEE. 2007. pp. 421-430.
12. Ernst M. D. Static and dynamic analysis: Synergy and duality. WODA 2003: ICSE Workshop on Dynamic Analysis. 2003. pp. 24-27.
13. Linn C., Debray S. Obfuscation of executable code to improve resistance to static disassembly. Proceedings of the 10th ACM conference on Computer and communications security. ACM, 2003. pp. 290-299.
14. Крис К. Искусство дизассемблирования. СПб.: БХВ-Петербург. 2008, 896 с.

15. Алексеев В. Е., Таланов В. А. Графы и алгоритмы. Структуры данных. Модели вычислений. М.: Интернет-университет информационных технологий, 2006, 67 с.

REFERENCES:

[1] Aleinikov S. I., Bogatov A. Protecting programs against disassembling. Proceedings of Institute for system programming of the Russian Academy of Sciences. 2006. T. 11. Pp. 9-16. (in Russian).
[2] Collberg C., Thomborson C. and Low D. A Taxonomy of Obfucating Transformations,Technical Report N148, Department of Computer Science. The University of Auckland, July, 1997.
[3] Hlestov A. D. Nicholas K. Y., Obfuscation and protection methods software products. Vestnik UrFO. Security in the information sphere. 2015. No. 2. P. 7-10. (in Russian).
[4] Yarmolik V.N., Portyanko S.S. State of the ART in Software Ownership Protection. Computer Information Systems and Industrial Management Application Editors: Khalid Saeed, Romuald Mosdorf, Olli-Pekka Hilmola. Bialystok, Poland. 2003. pp. 188-195.
[5] Collberg C.S., Thomborson C. Watermarking, Tamper-proofing and Obfuscation – Tools for Software protection. Computer Science Technical Report #170. University of Auckland, Aucland, New Zealand. February 10, 2000.
[6] Kozubenko Y. V. copyright Protection for computer programs in the criminal, administrative and civil proceedings. Wolters Kluwer. Russia, 2009, 344 p. (in Russian).
[7] Alexander G. Gounares Fingerprinting Executable Code. Patent US20120317421 A1, date 19.07.2012.
[8] Anckaert B. et al. Steganography for executables and code transformation signatures. International Conference on Information Security and Cryptology. Springer Berlin Heidelberg. 2004. pp. 425-439.
[9] Zaidan A. A. et al. Novel approach for high (secure and rate) data hidden within triplex space for executable file. Scientific Research and Essays. 2010. vol. 5. no. 15. pp. 1965-1977.
[10] Collberg C., Thomborson C. Software watermarking: Models and dynamic embeddings. Proceedings of the 26th ACM SIGPLAN-SIGACT symposium on Principles of programming languages. ACM. 1999. pp. 311-324.
[11] Moser A., Kruegel C., Kirda E. Limits of static analysis for malware detection. Computer security applications conference, 2007. ACSAC 2007. Twenty-third annual. IEEE. 2007. pp. 421-430.
[12] Ernst M. D. Static and dynamic analysis: Synergy and duality. WODA 2003: ICSE Workshop on Dynamic Analysis. 2003. pp. 24-27.
[13] Linn C., Debray S. Obfuscation of executable code to improve resistance to static disassembly. Proceedings of the 10th ACM conference on Computer and communications security. ACM, 2003. pp. 290-299.
[14] Chris K. the Art of disassembly. SPb.: BHV-Petersburg. 2008, 896 p. (in Russian).
[15] Alekseev V. E., Talanov V. A. Graphs and algorithms. The structure of the data. The model calculations. M.: Internet-University of information technologies, 2006, 67 p. (in Russian).

Поступила в редакцию – 20 февраля июля 2017 г. Окончательный вариант – 21 мая 2017 г.
Received – February 20, 2017. The final version – May 21, 2017.

Алина В. Трепачева
Южный федеральный университет,
ул. Большая Садовая, 105/42, г. Ростов-на-Дону, 344006, Россия,
e-mail: alina1989malina@ya.ru, ORCID iD 0000-0001-7625-569X

О СООТНОШЕНИЯХ МЕЖДУ АТАКАМИ НА СИММЕТРИЧНЫЕ ШИФРЫ, ГОМОМОРФНЫЕ НАД КОЛЬЦОМ ВЫЧЕТОВ[1]

DOI: http://dx.doi.org/10.26583/bit.2017.2.09

Аннотация. Работа посвящена изучению криптостойкости симметричных гомоморфных шифров над кольцами вычетов. Основной задачей является выяснить, возможно ли установить эквивалентность между атакой только по шифртекстам (АТШ) и атакой по известным открытым текстам (АИОТ) для таких шифров. Для этого вводится понятие сводимости между атаками и дается достаточное условие сводимости от АТШ к АИОТ. Основная идея заключается в том, что для доказательства сводимости от АТШ к АИОТ необходимо найти функцию над кольцом вычетов, которая является эффективно вычислимой и имеет небольшой размер образа по сравнению с размером всего кольца вычетов. Исследование наличия сводимости интересно тем, что оно может позволить лучше понять уровень криптостойкости существующих симметричных гомоморфных криптосистем (ГК). Поскольку для большинства из них в литературе уже установлена уязвимость к АИОТ, то доказательство существования сводимости может показать, что эти криптосистемы не стойки даже к АТШ, а, следовательно, совсем не стойки и непригодны для применения. Приводится пример сводимости АТШ к АИОТ для случая, когда кольцо вычетов является простым полем. На основе этого примера описывается эффективная АТШ на одну симметричную ГК в случае кольца вычетов небольшого размера. Также отдельно рассматривается случай, когда кольцо вычетов для ГК строится по труднофакторизуемому модулю n. Для таких n на данный момент не известен эффективный алгоритм построения эффективно вычислимых функций с «небольшим» образом. Ввиду этого дальнейшая работа по криптоанализу существующих симметричных ГК будет направлена на изучение свойств функций над кольцами вычетов по труднофакторизуемым модулям.

Ключевые слова: гомоморфное шифрование над кольцом вычетов, полностью гомоморфное шифрование, задача RSA, задача факторизации чисел, атака только по шифртекстам, атака по известным открытым текстам.

Для цитирования. ТРЕПАЧЕВА, Алина В. О соотношениях между атаками на симметричные шифры, гомоморфные над кольцом вычетов. Безопасность информационных технологий, [S.l.], v. 24, n. 2, p. 82-91, june 2017. ISSN 2074-7136. Доступно на: <https://bit.mephi.ru/index.php/bit/article/view/108>. Дата доступа: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.09.

[1]***Благодарности:*** Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 15-07-00597 А

Alina V. Trepacheva
Southern Federal University,
Bolshaya Sadovaya St., 105/42 Rostov-on-Don, 344006, Russia
e-mail: alina1989malina@ya.ru, ORCID iD 0000-0001-7625-569X

On the relations between the attacks on symmetric homomorphic encryption over the residue ring[1]

DOI: http://dx.doi.org/10.26583/bit.2017.2.09

Abstract. The paper considers the security of symmetric homomorphic cryptosystems (HC) over the residue ring. The main task is to establish an equivalence between ciphertexts only attack (COA) and known plaintexts attack (KPA) for HC. The notion of reducibility between attacks

and sufficient condition of reducibility from COA to KPA are given for this purpose. The main idea is: to prove reducibility from COA to KPA we need to find a function over residue ring being efficiently computable and having a small image size comparing with the size of residue ring. The study of reducibility existence is important since it allows to understand better the security level of symmetric HC proposed in literature. A vulnerability against KPA has been already found for the majority of these HC. Thus the reducibility presence can demonstrate that cryptosystems under the study are not secure even against COA, and therefore they are totally insecure and shouldn't be used in practice. We give an example of reducibility from COA to KPA for residue ring being a simple field. Based on this example we show an efficient COA on one symmetric HC for small field. Also we separately consider the case of residue ring composed using number n being hard-to-factor. For such n an efficient algorithm to construct an efficiently computable function with small image is unknown so far. So further work related to cryptanalysis of existing symmetric HC will be directed into study of functions properties over residue rings modulo numbers hard for factorization.

Keywords: ring-homomorphic encryption, fully homomorphic encryption, RSA problem, number factorization problem, ciphertexts only attack, known plaintext attack.

For citation. TREPACHEVA, Alina V. On the Relations between the Attacks on Symmetric Homomorphic Encryption over the Residue Ring. IT Security (Russia), [S.l.], v. 24, n. 2, p. 82-91, june 2017. ISSN 2074-7136. Available at: <https://bit.mephi.ru/index.php/bit/article/view/108>. Date accessed: 23 june 2017. doi:http://dx.doi.org/10.26583/bit.2017.2.09.

Acknowledgements: This work is executed at financial support RFBR, research project No. 15-07-00597 A.

Введение

Эффективное и криптостойкое полностью гомоморфное шифрование (ПГШ) дает принципиально новые возможности по обеспечению информационной безопасности в таких областях как облачные вычисления, обработка медицинских и финансовых данных, поскольку позволяет проводить любые вычисления над данными в зашифрованном виде без знания ключа расшифрования в недоверенной среде [1]. Именно поэтому интерес к ПГШ в последние годы все больше возрастает. Предложено уже большое количество полностью гомоморфных криптосистем (ПГК). Одним из наиболее популярных направлений в этой области является построение ПГК с открытым ключом [2,3], основанных на теории решеток. Родоначальником этого направления была ПГК из работы [3], разработанная Джентри в 2009 году. Достоинством этих ПГК является строго доказуемая криптостойкость против атаки по выбранным открытым текстам (АВОТ). Однако на практике их применение на данный момент весьма ограничено из-за того, что вычисления над зашифрованными данными сильно усложняются по сравнению с вычислениями над исходными данными [1].

Вследствие этого было предложено большое количество альтернативных вариантов симметричных гомоморфных криптосистем (ГК) над кольцами вычетов [4-15]. Особенно активны попытки построить симметричные ПГК, криптостойкость которых основывалась бы некоторым образом на задаче факторизации больших чисел [5-14], являющейся эталоном сложной задачи в компьютерной безопасности [16]. Некоторые из ГК [4-15] даже были использованы в реальных приложениях. К примеру, авторы [5] использовали свою ПГК для защиты облачной БД [17], а ГК из работ [12,15] были использованы для защиты данных в беспроводных сенсорных сетях [18,19].

Однако строгое обоснование защищенности ни для одной ГК из работ [4-15] не было приведено. Проведенный в работах [20-29] криптоанализ показал нестойкость ГК [4-13,15] к атаке по известным открытым текстам (АИОТ). Как следствие, к примеру, в работе [30] было выявлено, что защищенная БД из [17] имеет уязвимость.

Но вопрос о криптостойкости относительно атаки только по шифртекстам (АТШ) не был изучен в полной мере. Целью данной работы является предложить метод анализа

криптостойкости к АТШ указанных симметричных криптосистем, гомоморфных над кольцами вычетов. Основная идея этого метода – попробовать свести АТШ к АИОТ.

Вводные замечания и обозначения

Определение 1. Гомоморфной над кольцом криптосистемой (ГКК, англ. ring-homomorphic encryption) называют криптосистему, у которой множество открытых текстов P имеет алгебраическую структуру кольца и обладающую свойствами:

$$D\big(E(m_1)\oplus E(m_2)\big) = m_1 + m_2, \quad (1)$$
$$D\big(E(m_1)\odot E(m_2)\big) = m_1 \cdot m_2, \quad (2)$$

где E – функция шифрования, D – функция расшифрования, $\oplus,\odot$ – операции над кольцом шифртекстов C, $+$, $\cdot$ – операции над кольцом открытых текстов P.

Гомоморфный шифр называется *компактным*, если размер шифртекста, получающегося в результате проведения вычислений, ограничен некоторым числом, полиномиальным от параметров шифра.

ПГК является компактной гомоморфной ГКК, а также должно выполняться, что $\oplus,\odot$ можно вычислить над шифртекстами любое число раз корректно и эффективно.

В данной работе будут рассматриваться ГКК, для которых $P = \mathbb{Z}_n, n \in \mathbb{N}, n > 1$.

Определение 2. Эффективно вычислимой над $\mathbb{Z}_n$ функцией будем называть функцию, которую можно вычислить за $O(\log n)$ операций $+$, $\cdot$ кольца $\mathbb{Z}_n$.

Простейшими примером эффективно вычислимой функций над $\mathbb{Z}_n$ является полином $(x + c)^n$, т.к. существуют алгоритмы, которые могут возвести в степень n за $\log_2 n$ шагов (например, т.н. алгоритмы двоичного возведения в степень [31]).

Будем обозначать образ функции $f(x)$ на множестве S как $\mathrm{Im}(f(S))$, а количество элементов в этом образе как $|\mathrm{Im}(f(S))|$.

Основные типы атак на криптосистемы и определение сводимости

Известно, что существуют три основных типа атак на криптосистемы [32]:

1) АТШ (ciphertexts-only attack), подразумевающая что криптоаналитику доступны только шифртексты, изготовленные на одном ключе;

2) АИОТ (known-plaintexts attack), в которой криптоаналитик перехватил пары $(m_1, E(m_1)), \dots, (m_s, E(m_s))$, изготовленные на одном ключе;

3) АВОТ (chosen-plaintexts attack), предполагающая что криптоаналитик имеет доступ к шифрующему устройству и может зашифровать выбранные по своему усмотрению открытые тексты, т.е. может получить пары $(m_1, E(m_1)), \dots, (m_s, E(m_s))$ для выбранных m_i.

Есть также и другие виды атак, но здесь мы опускаем их описание.

Для криптосистем с открытым ключом наиболее актуально обоснование криптостойкости относительно АВОТ, поскольку получение пар «открытый текст – шифртекст» общедоступно. Поэтому для ПГК с открытым ключом [2,3] обычно доказывают АВОТ-стойкость. Для симметричных же криптосистем имеет смысл рассматривать стойкость против АТШ и АИОТ отдельно. В частности, анализируемые в данной работе ГКК [4-15], как уже упоминалось, симметричны.

Атаку на криптосистему можно рассматривать как алгоритмическую задачу [33] и в этом случае можно рассмотреть вопрос о сводимости одной атаки к другой.

Определение 3. Сводимость атаки $\mathcal{A}$ к атаке $\mathcal{B}$ – это эффективно вычислимая функция, которая преобразует исходные данные для атаки $\mathcal{A}$ в исходные данные для атаки $\mathcal{B}$.

Для любой криптосистемы есть тривиальные сводимости: АИОТ → АТШ, АВОТ → АИОТ. Однако в случае гомоморфных шифров могут иметь место и обратные сводимости, которые сделают некоторые атаки эквивалентными для ГШ. В частности, исследование наличия сводимости АТШ → АИОТ для ранее упомянутых ГКК [4-15] может позволить ответить на вопрос об их защищенности относительно АТШ.

Пример сводимости АТШ к АИОТ для случая, когда пространство открытых текстов – кольцо вычетов по модулю простого числа

Проанализируем сводимость АТШ к АИОТ на ГКК для случая $P = \mathbb{Z}_p$, где p – простое число. Как известно, по теореме Эйлера выполняется $\forall m \in \mathbb{Z}_p: m^p - m = 0(mod\ p)$. Поэтому, пользуясь гомоморфными свойствами шифра, криптоаналитик может составить функцию $f(x) = x^p - x$, благодаря которой он может превратить АТШ в АИОТ. Для этого ему достаточно подставить любой шифртекст в f и он получит шифртекст нуля.

Рассмотрим этот процесс более подробно на примере ГКК [4], для которой $P = \mathbb{Z}_p$. Шифрование в [4] сводится к формуле $c(x) = m + k(x) \cdot r(x) \in \mathbb{Z}_p[x]$, где $m \in \mathbb{Z}_p$ – открытый текст, $k(x) \in \mathbb{Z}_p[x]$ – секретный ключ, $r(x) \in \mathbb{Z}_p[x]$ – случайный полином, полученный по равномерному распределению, $\deg(k), \deg(r) > 0$. Основная идея АИОТ на [4], описанной в [22,23], следующая: пусть криптоаналитик перехватил пары $(m_1, c_1(x)), \dots, (m_s, c_s(x))$, зашифрованные на $k(x)$; тогда для раскрытия $k(x)$ он может вычислить $g(x) = \text{НОД}(c_1(x) - m_1, \dots, c_s(x) - m_s)$. В силу случайности $r_i(x)$ полином $g(x)$ с вероятностью ≈ 1 равен $k(x)$ для небольших значений s. К примеру, для $\mathbb{Z}_2$ достаточно взять $s > 5$.

На число p в [4] не накладывается ограничений. Рассмотрим случай, когда p – простое. Предположим криптоаналитик имеет $c_1(x), \dots, c_s(x)$, изготовленные на $k(x)$, и знает p. Он может вычислить $c_i'(x) = (c_i(x))^p - c_i(x),\ i = 1, \dots, s$ и это даст ему пары для АИОТ $(0, c_1'(x)), \dots, (0, c_s'(x))$.

Однако необходимо отметить, что вероятностное распределение шифртекстов $c_i'(x)$ отлично от того, по которому получаются «свежие» шифртексты (произведенные непосредственно алгоритмом зашифрования). Имеем

$$c_i'(x) = (c_i(x))^p - c_i(x) = k(x) \cdot r_i(x) \cdot ((k(x) \cdot r_i(x))^{p-1} - 1).$$

Полином данного вида будет нацело делиться на $x^p - x$ для любых $k(x), r_i(x)$. Однако, для «свежих» шифртекстов такое свойство выполняться не будет.

АИОТ из [22,23], описанная ранее, была рассчитана именно на «свежие» шифртексты. В данном же случае в неё необходимо внести поправку. Полином $g(x)$ необходимо разделить на $x^p - x$, т.е. провести нормировку $g(x)$. В результате мы получим $k(x)$, если $k(x)$ не имеет линейных сомножителей. Если же имеет, то мы получим некоторый сомножитель $k(x)$, но, очевидно, и его достаточно для взлома ГКК [4]. Описанные выше действия формализованы в виде *Алгоритма 1*.

Алгоритм 1 CoaAdHoc($c_1, \dots, c_t$)

Вход: «свежие» шифртексты $c_1(x), \dots, c_s(x)$ ГКК [4], простое число p
Выход: $k(x)$ или его множитель
1: **for** *i=1* to *s* **do**
2: $c_i^*(x) := (c_i(x))^p - c_i(x)$
3: **end for**
4: **return** НОД$(c_1^*(x), \dots, c_s^*(x))/(x^p - x)$

Рассмотрим теперь вопрос об эффективности алгоритма 1. Функция $f(x) = x^p - x$ является эффективно вычислимой над $\mathbb{Z}_p$ для любого p. Однако описанная ГКК некомпактна, т.к. при умножении размеры шифртекстов растут с экспоненциальной скоростью. Поэтому для больших p эта атака не будет эффективна. Для того чтобы получить эффективную атаку для любого p необходима функция f, имеющая небольшую степень, к примеру $log_2 p$.

В случае, когда $P = \mathbb{Z}_n$ и n – составное число, можно использовать функцию $f(x) = x^{\varphi(n)} - x$, где $\varphi(n)$ – функция Эйлера. Если n – небольшое легко факторизуемое число, то мы снова получаем эффективную атаку на ГКК [4]. Если же n трудно факторизуемо, то так действовать не получится, поскольку вычисление $\varphi(n)$ будет трудной задачей.

Обобщение идеи о сводимости АТШ к АИОТ

Итак, как видно из примера в предыдущем разделе, для получения сводимости АТШ → АИОТ была использована функция $f: P \to P$, которая отображает любой открытый текст $m \in P$ в одно фиксированное значение, т.е. $|\mathrm{Im}(f(P))| = 1$. *Однако, в общем, для получения сводимости атаки может быть достаточно, чтобы выполнялось соотношение* $|Im(f(P))| \ll |P|$. *И тогда полный перебор элементов* $Im(f)$ *не затруднит атаку.*

Определение 4. Назовем функцию $f: P \to P$ σ - *логарифмически сжимающей*, если $|\mathrm{Im}(f(P))| < \sigma \cdot \log_2 |P|$, т.е. количество элементов в её образе не превышает двоичного логарифма от общего количества элементов в P, умноженного на константу $\sigma \in \mathbb{R}$, где $\sigma \ll |P|$

В качестве примера приведем σ - логарифмически сжимающую функцию над $\mathbb{Z}_6$ -- полином $f(x) = x^2 + x + 1$ переводит элементы $\mathbb{Z}_6$ в два элемента: 1 и 3.

Также для наших целей подойдет не только функция «гарантированно» отправляющая все элементы P в «небольшой» образ, но и функция, которая «с большой вероятностью» отправляет их туда.

Определение 5. Назовем функцию $f: P \to P$ Ω - *вероятностно* σ - *логарифмически сжимающей*, если существует такое подмножество элементов $S \subseteq P$, $|S| > \Omega \cdot |P|$ что $|\mathrm{Im}(f(S))| < \sigma \cdot \log_2 |P|$, $\sigma, \Omega \in \mathbb{R}$, $0 < \Omega \leq 1$.

Заметим, что σ – логарифмически сжимающая функция является частным случаем P Ω - вероятностно σ - логарифмически сжимающей для $\Omega = 1$.

Основные теоремы о достаточном условии эквивалентности АИОТ и АТШ для АГК сформулируем уже для $P = \mathbb{Z}_n$, $n \in \mathbb{N}, n > 1$.

Теорема 1. Если существует алгоритм, работающий на всех входах за не более чем $O(\log_2 n)$ шагов, который для заданного пространства открытых текстов $\mathbb{Z}_n$ некоторой ГКК, возможно не обладающей свойством компактности, и некоторого σ выдает Ω-

вероятностно σ - логарифмически сжимающую функцию $f: \mathbb{Z}_n \to \mathbb{Z}_n$ с $\Omega \approx 1$, выражаемую полиномом степени не более $O(\log_2 n)$, то АТШ и АИОТ на эту ГКК эквивалентны.

Теорема 2. Если существует алгоритм, работающий на всех входах за не более чем $O(\log n)$ шагов, который для заданного пространства открытых текстов $\mathbb{Z}_n$ некоторой ПГК, и некоторого σ выдает эффективно вычислимую Ω-вероятностно σ-логарифмически сжимающую функцию $f: \mathbb{Z}_n \to \mathbb{Z}_n$ с $\Omega \approx 1$, то АТШ и АИОТ на эту ГКК эквивалентны.

Для иллюстрации теорем приведен общий алгоритм $\mathrm{CoaHom}(c_1, \dots, c_s)$ для осуществления АТШ на ГКК посредством сведения к АИОТ с применением σ - логарифмически сжимающей функции $f: \mathbb{Z}_n \to \mathbb{Z}_n$.

Алгоритм 2 $\mathrm{CoaHom}(c_1, \dots, c_s)$

Вход: шифртексты $c_1, \dots, c_s$ криптосистемы, произведенные на одном ключе; s – число пар «открытый текст, шифртекст», необходимых для АИОТ; функция f

Выход: секретный ключ sk в случае успешной атаки

1: **for** *i=1* to *s* **do**
2: $c_i^* := f(c_i)$
3: **end for**
4: **for** $(p_1, \dots, p_s) \in Im(f) \times \dots \times Im(f)$ **do**
5: sk′ = АИОТ $((p_1, c_1^*), \dots, (p_s, c_s^*))$
5: **if** АИОТ $((p_1, c_1^*), \dots, (p_s, c_s^*))$ прошла успешно **then**
6: sk = Normalize(sk′)
7: **return** sk
8: **end if**
9: **end for**

Функция Normalize, упомянутая в алгоритме, необходима для нормализации результата АИОТ атаки. Она может быть нужна, если в результате применения функции f к шифртекстам поменяются их вероятностные распределения из-за чего у них появятся дополнительные свойства и закономерности. В общем случае функция Normalize будет существенно зависеть от вида функции f и самой ГКК. Для примера, рассмотренного в предыдущем разделе, Normalize реализовывалась делением результата АИОТ на полином $x^p - x$.

тметим также, что критерий того, что АИОТ прошла успешна в общем случае тоже сформулировать затруднительно, т.к. он зависит от ГКК и, возможно, самой АИОТ. В качестве примера рассмотрим случай, когда $P = \mathbb{Z}_n$, где n – труднофакторизуемый RSA-модуль. Опираясь на существующий опыт по проведению АИОТ на ГКК с таким P [20-25], можно сказать, что в этом случае критерием успеха атаки может, к примеру, служить раскрытие факторизации числа n.

Напоследок отметим, что общая асимптотическая сложность алгоритма $CoaHom(c_1, \dots, c_s)$ составляет $\approx O(\log_2^s n) \cdot Complexity_{\text{АИОТ}}$.

О сводимости АТШ к АИОТ для случая, когда пространство открытых текстов – кольцо вычетов по модулю труднофакторизуемого числа

Большая часть существующих симметричных ГКК [5-14], как уже отмечалось, построена таким образом, что их защищенность пытаются обосновать с привлечением классической и хорошо изученной трудной задачи факторизации больших чисел. Эти ГКК взламываются при факторизации некоторого числа n, входящего в их конструкцию. Хотя

строгая сводимость криптостойкости к этой задаче ни в одной из указанных работ не была установлена и факторизация n не является необходимым условием для их взлома (по крайней мере для АИОТ это уже известно точно согласно работам по их криптоанализу [20-27]).

Однако актуальность построения доказуемо стойкого и эффективного полностью гомоморфного шифра на основе задачи факторизации по-прежнему не теряет смысла и актуальности, поскольку в работе [34] было предъявлено доказательство того, что теоретически возможно построить АТШ-криптостойкую ГКК над $\mathbb{Z}_n$, основанную на задаче факторизации n.

Существующие ГКК, связанные с задачей факторизации чисел (ФГГК), можно условно классифицировать по структуре их пространств открытых текстов на 2 группы:

1) $P = \mathbb{Z}_n$, где $n = p \cdot q$ – труднофакторизуемое число, p, q – большие криптографически стойкие, простые числа [6-14];
2) $P = \mathbb{Z}_p$, однако $n = p \cdot q$ также используется в алгоритме шифрования [5].

Для ФГКК 2-го типа из работы [5] была показана её нестойкость к АИОТ в [20]. Однако, применение описанной выше концепции для анализа её стойкости к АТШ выглядит на данный момент затруднительно, поскольку фактически авторы [5] скрывают P, нарушая тем самым принцип Керхгоффса. И пока не вполне ясно, как можно было бы построить необходимую для сводимости сжимающую функцию $f: P \to P$ в предположении, что криптоаналитику не известно P.

Для ФГКК 1-го типа σ - логарифмически сжимающую функцию можно построить, например, с помощью интерполяции. Однако, возможно, что степень полинома, реализующего эту функцию, будет слишком большой и как следствие вычисление функции будет неэффективно. Можно также действовать и полным перебором, однако это неэффективно. Пытаться получить такую функцию, случайно выбрав некоторый полином, также представляется неэффективным подходом, хотя данный вопрос требует отдельного подробного изучения (в частности, проведение компьютерных экспериментов для небольших n с вычислением процента сжимающих функций над $\mathbb{Z}_n$ по сравнению с общим числом функций над $\mathbb{Z}_n$). Итого, вопрос об эффективном алгоритме построения эффективно вычислимых логарифмически сжимающих функций для произвольного труднофакторизуемого числа n на данный момент является открытым и требует дальнейшего изучения.

Заключение

Проведено исследование, устанавливающее взаимосвязь между атаками только по шифртекстам и по известным открытым текстам на шифры, гомоморфные над кольцом вычетов $\mathbb{Z}_n$. Установлено, что вопрос об эквивалентности этих атак можно свести к поиску семейства сжимающих функций над $\mathbb{Z}_n$, а также эффективного алгоритма, который для каждого $\mathbb{Z}_n$ предъявляет такую функцию.

В случае построения эффективного алгоритма нахождения сжимающих функций над $\mathbb{Z}_n$, где $n = p \cdot q$ – труднофакторизуемое число, можно будет сделать окончательные выводы о криптостойкости гомоморфных шифров из работ [6-13] к атаке только по шифртекстам. Поэтому дальнейшая работа по криптоанализу этих ГКК будет направлена на изучение свойств функций над $\mathbb{Z}_n$.

СПИСОК ЛИТЕРАТУРЫ:

1. Бабенко Л. К., Буртыка Ф. Б., Макаревич О. Б., Трепачева А. В. Полностью гомоморфное шифрование (обзор). Вопросы защиты информации. 2016, № 3, С. 3–25.
2. Gentry, Craig. "Fully homomorphic encryption using ideal lattices." STOC. Vol. 9. No. 2009. 2009.

3. Vaikuntanathan, V. Computing blindfolded: New developments in fully homomorphic encryption. Proceedings of IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS), 2011.
4. Жиров А. О., Жирова О. В., Кренделев С. Ф. Безопасные облачные вычисления с помощью гомоморфной криптографии. Безопасность информационных технологий, № 1, 2013. С. 6-12.
5. Zhirov A., Zhirova O, Krendelev S. Practical fully homomorphic encryption over polynomial quotient ring. Proceedings of World congress on internet security (WorldCIS), IEEE. 2013, С. 70-75. DOI: 10.1109/WorldCIS.2013.6751020
6. Ростовцев А., Богданов А., Михайлов М. Метод безопасного вычисления полинома в недоверенной среде с помощью гомоморфизмов колец. Проблемы информационной безопасности. Компьютерные системы. 2011. № 2. С. 76-85.
7. Yagisawa M. Fully Homomorphic Encryption with Composite Number Modulus. IACR Cryptology ePrint Archive. 2015. № 1040. URL: https://eprint.iacr.org/2015/1040.pdf (дата обращения: 25.01.2017).
8. Kipnis A., Hibshoosh E. Efficient Methods for Practical Fully Homomorphic Symmetric-key Encrypton, Randomization and Verification. IACR Cryptology ePrint Archive. 2012. № 637. URL: https://eprint.iacr.org/2012/637.pdf (дата обращения: 25.01.2017).
9. Xiao L., Bastani O., Yen I. L. An Efficient Homomorphic Encryption Protocol for Multi-User Systems. IACR Cryptology ePrint Archive. 2012. №. 193 URL: https://eprint.iacr.org/2012/193.pdf (дата обращения: 22.01.2017).
10. Gupta C. P., Sharma I. Department of Computer Sciences and Engineering Rajasthan Technical University, Kota, India. Network of the Future (NOF), 2013 Fourth International Conference on the. – IEEE, 2013. С. 1-4.
11. Gupta C. P. Fully Homomorphic Encryption Scheme with Symmetric Keys : дис. – Department of Computer Science & Engineering University College of Engineering, Rajasthan Technical University, Kota, 2013
12. J. Domingo-Ferrer, A new privacy homomorphism and applications, Information Processing Letters, vol. 60, no. 5, С. 277–282, 1996.
13. Chan A. C. F. Symmetric-key homomorphic encryption for encrypted data processing. Communications, 2009. ICC'09. IEEE International Conference on. – IEEE, 2009. – С. 1-5.
14. Gavin G. A general framework for building noise-free homomorphic cryptosystems. IACR Cryptology ePrint Archive. 2015. № 821. URL: https://eprint.iacr.org/2015/821.pdf (дата обращения: 25.01.2017).
15. J. Domingo-Ferrer, A provably secure additive and multiplicative privacy homomorphism. Information Security. 2002, С.471–483.
16. Rivest R. L., Kaliski Jr B. RSA problem. Encyclopedia of cryptography and security. – Springer US, 2011. – С. 1065-1069.
17. Shatilov K., Boiko V., Krendelev S., Anisutina D., Sumaneev A. Solution for secure private data storage in a cloud. Proceedings of Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2014, С. 885-889. DOI: 10.15439/2014F43
18. Ertaul L., Yang J. H. Implementation of Domingo-Ferrer's a new privacy homomorphism (df a new ph) in securing wireless sensor networks (wsn). Security and Management. Citeseer, 2008, С. 498–504.
19. Jariwala V., Jinwala D. Evaluating homomorphic encryption algorithms for privacy in wireless sensor networks. International Journal of Advancements in Computing Technology, vol. 3, no. 6, 2011.
20. Трепачева А.В. О криптоанализе одной полностью гомоморфной криптосистемы на основе задачи факторизации. Безопасность информационных технологий. 2015, № 4, С. 19-25
21. Трепачева А.В. Улучшенная атака по известным открытым текстам на гомоморфную криптосистему Доминго-Феррера. Труды Института системного программирования Российской академии наук. 2014. Том 25, Вып. 5. С. 83-98. DOI: 10.15514/ISPRAS-2014-26(5)-4
22. Трепачева А. В. Криптоанализ шифров, основанных на гомоморфизмах полиномиальных колец. Известия Южного федерального университета. Технические науки, том 157, №. 8, С. 96-107, 2014.
23. Trepacheva A., Babenko L. Known plaintexts attack on polynomial based homomorphic encryption. Proceedings of the 7th International Conference on Security of Information and Networks. – ACM, 2014. – С. 157.
24. Трепачева А.В. Криптоанализ симметричных полностью гомоморфных линейных криптосистем на основе задачи факторизации чисел. Известия Южного федерального университета. Технические науки № 5 (том 166) (2015).

25. Трепачева А.В. Криптоанализ полностью гомоморфных криптосистем, основанных на алгебре октонионов. Обозрение прикладной и промышленной математики 23(4) 2016
26. Vizár D., Vaudenay S. Analysis of Chosen Symmetric Homomorphic Schemes. Central European Crypto Conference. – 2014. – №. EPFL-CONF-198992.
27. Tsaban B., Lifshitz N. Cryptanalysis of the MORE symmetric key fully homomorphic encryption scheme. Journal of Mathematical Cryptology. – 2014.
28. Cheon J. H., Kim W.-H., Nam H. S. Known-plaintext cryptanalysis of the Domingo-Ferrer algebraic privacy homomorphism scheme. Information Processing Letters, vol. 97, no. 3, С. 118–123, 2006.
29. Wagner, Cryptanalysis of an algebraic privacy homomorphism. Information Security. Springer, 2003, С. 234–239.
30. Бабенко Л.К., Трепачева А.В. Анализ защищенности одной криптографической системы для защиты конфиденциальности данных в облачных вычислениях. Безопасность информационных технологий. 2016, № 1, С. 11-15
31. Gordon D. M. A survey of fast exponentiation methods. Journal of algorithms. – 1998. – Т. 27. – №. 1. – С. 129-146.
32. Goldreich, O. Foundations of Cryptography–A Primer. Foundations and Trends® in Theoretical Computer Science 1.1 (2005): 1-116.
33. Гэри, М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. Москва: Мир, 1982. С. 416.
34. Altmann K., Jager T., Rupp A. On black-box ring extraction and integer factorization. International Colloquium on Automata, Languages, and Programming. 2008. С. 437-448. DOI : 10.1007/978-3-540-70583-3_36.

REFERENCES:

[1] Babenko L. K., Burtica F. B., Makarevich O. B., Trubacheva A. V. Fully homomorphic encryption (review). Problems of information security. 2016, no. 3, Pp. 3-25. (in Russian).
[2] Gentry, Craig. "Fully homomorphic encryption using ideal lattices." STOC. Vol. 9. No. 2009. 2009.
[3] Vaikuntanathan, Vinod. "Computing blindfolded: New developments in fully homomorphic encryption." Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on. IEEE, 2011.
[4] A. O. Zhirov, O. V. Zhirova, and S. F. Krendelev. Secure cloud computing using homomorphic cryptography. Security of information technologies, № 1, 2013. С. 6-12. (in Russian)
[5] Zhirov A., Zhirova O, Krendelev S. Practical fully homomorphic encryption over polynomial quotient ring. World congress on internet security (WorldCIS), IEEE. 2013, С. 70-75. DOI: 10.1109/WorldCIS.2013.6751020
[6] A. B. Alexander Rostovtsev and M. Mikhaylov. Secure evaluation of polynomial using privacy ring homomorphisms. Cryptology ePrint Archive, Report 2011/024, 2011. http://eprint.iacr.org/
[7] Yagisawa M. Fully Homomorphic Encryption with Composite Number Modulus. IACR Cryptology ePrint Archive. 2015. № 1040. URL: https://eprint.iacr.org/2015/1040.pdf (дата обращения: 25.01.2017).
[8] Kipnis A., Hibshoosh E. Efficient Methods for Practical Fully Homomorphic Symmetric-key Encrypton, Randomization and Verification. IACR Cryptology ePrint Archive. 2012. № 637. URL: https://eprint.iacr.org/2012/637.pdf (дата обращения: 25.01.2017).
[9] Xiao L., Bastani O., Yen I. L. An Efficient Homomorphic Encryption Protocol for Multi-User Systems. IACR Cryptology ePrint Archive. – 2012. – №. 193
[10] Gupta C. P., Sharma I. Department of Computer Sciences and Engineering Rajasthan Technical University, Kota, India. Network of the Future (NOF), 2013 Fourth International Conference on the. – IEEE, 2013. – С. 1-4.
[11] Gupta C. P. Fully Homomorphic Encryption Scheme with Symmetric Keys : дис. – Department of Computer Science & Engineering University College of Engineering, Rajasthan Technical University, Kota, 2013
[12] J. D. i. Ferrer, "A new privacy homomorphism and applications," Information Processing Letters, vol. 60, no. 5, pp. 277–282, 1996.
[13] Chan A. C. F. Symmetric-key homomorphic encryption for encrypted data processing. Communications, 2009. ICC'09. IEEE International Conference on. – IEEE, 2009. – P. 1-5.

[14] Gavin G. A general framework for building noise-free homomorphic cryptosystems. IACR Cryptology ePrint Archive. 2015. № 821. URL: https://eprint.iacr.org/2015/821.pdf (дата обращения: 25.01.2017).

[15] J. Domingo-Ferrer, "A provably secure additive and multiplicative privacy homomorphism*," in Information Security. Springer, 2002, pp.471–483.

[16] Rivest R. L., Kaliski Jr B. RSA problem. Encyclopedia of cryptography and security. – Springer US, 2011. – С. 1065-1069.

[17] Shatilov K., Boiko V., Krendelev S., Anisutina D., Sumaneev A. Solution for secure private data storage in a cloud. Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2014, С. 885-889. DOI: 10.15439/2014F43.

[18] L. Ertaul and J. H. Yang, "Implementation of domingo ferrer's a new privacy homomorphism (df a new ph) in securing wireless sensor networks (wsn)." in Security and Management. Citeseer, 2008, pp. 498–504.

[19] V. Jariwala and D. Jinwala, "Evaluating homomorphic encryption algorithms for privacy in wireless sensor networks," International Journal of Advancements in Computing Technology, vol. 3, no. 6, 2011.

[20] Trepacheva A.V. On the cryptanalysis of a fully homomorphic cryptosystem based on the problem of factorization. Security of information technologies. 2015, № 4, p. 19-25 (in Russian).

[21] Trepacheva A.V. Improved attack known open texts on the homomorphic cryptosystem, Domingo-Ferrer.Proceedings of Institute for system programming of the Russian Academy of Sciences. 2014. Volume 25, Issue. 5. S. 83-98. DOI: 10.15514/ISPRAS-2014-26(5)-4 (in Russian).

[22] Trepacheva A. V. Cryptanalysis of ciphers based on the homomorphisms of polynomial rings. News of southern Federal University. Engineering science, volume 157, No.. 8, pp. 96-107, 2014. (in Russian).

[23] Trepacheva A., Babenko L. Known plaintexts attack on polynomial based homomorphic encryption. Proceedings of the 7th International Conference on Security of Information and Networks. – ACM, 2014. – С. 157.

[24] Trepacheva A.V. Cryptanalysis of symmetric linear fully homomorphic public key cryptosystems based on the problem of factorization of numbers. News of southern Federal University. Technical Sciences, No. 5 (volume 166) (2015). (in Russian).

[25] Trepacheva A.V. Cryptanalysis of a fully homomorphic public key cryptosystems based on algebra of octonions. Review of applied and industrial mathematics 23(4) 2016 (in Russian).

[26] Vizár D., Vaudenay S. Analysis of Chosen Symmetric Homomorphic Schemes. Central European Crypto Conference. – 2014. – №. EPFL-CONF-198992.

[27] Tsaban B., Lifshitz N. Cryptanalysis of the MORE symmetric key fully homomorphic encryption scheme. Journal of Mathematical Cryptology. – 2014.

[28] J. H. Cheon, W.-H. Kim, and H. S. Nam, "Known-plaintext cryptanalysis of the domingo-ferrer algebraic privacy homomorphism scheme," Information Processing Letters, vol. 97, no. 3, pp. 118–123, 2006.

[29] Wagner, "Cryptanalysis of an algebraic privacy homomorphism," inInformation Security. Springer, 2003, pp. 234–239.

[30] Babenko L.K., Trepacheva A.V. Security analysis of a single cryptographic system to protect data privacy in cloud computing. Security of information technologies. 2016, no. 1, Pp. 11-15 (in Russian).

[31] Gordon D. M. A survey of fast exponentiation methods. Journal of algorithms. – 1998. – Т. 27. – №. 1. – С. 129-146.

[32] Goldreich, Oded. "Foundations of Cryptography–A Primer." Foundations and Trends® in Theoretical Computer Science 1.1 (2005): 1-116.

[33] Gari, M., and D. Dzhonson. Computers and trudnoreshaemyh tasks. Moscow: Mir, 1982. S. 416. (in Russian).

[34] Altmann K., Jager T., Rupp A. On black-box ring extraction and integer factorization. International Colloquium on Automata, Languages, and Programming. 2008. С. 437-448. DOI : 10.1007/978-3-540-70583-3_36

Поступила в редакцию – 22 февраля июля 2017 г. Окончательный вариант – 23 мая 2017 г.
Received – February 22, 2017. The final version – May 23, 2017.

АННОТАЦИИ

Александр М. Алюшин, Сергей В. Дворянкин

Использование речевых технологий для защиты документооборота

Ключевые слова: спектрограмма, аудиомаркер, распознавание, документация, графическая картинка

Несмотря на то, что мы живем в веке информационных технологий, бумажный документооборот не теряет своей актуальности. Выделены основные аспекты защиты документов, связанные с защитой их контекстной и юридической составляющих. К контекстной составляющей отнесен смысловой информационный аспект документа. К юридической составляющей отнесены факты и условия создания, утверждения, согласования данного документа конкретными лицами. Показана актуальность защиты документов в связи с возможными террористическими угрозами. Показана важность для эффективной защиты документов такого фактора как время выявления фальсификации. Представлен анализ требований к данному фактору для документов различного характера – финансовых, юридических, управляющих. Выделены управляющие документы, использующиеся для оперативного управления опасными объектами. Показано, что их умышленная фальсификация может привести к возникновению аварий и катастроф техногенного происхождения, а также человеческим жертвам. Дан сравнительный анализ применяющихся в настоящее время на практике методов защиты документов. Выделены биометрические и небиометрические методы защиты документов. Представлен анализ их недостатков. Сделан вывод о перспективности применения технологии защиты документов на основе речевой подписи. Рассмотрены основные этапы обработки голосовой информации при реализации данной технологии. Разработано программное обеспечение, реализующее новую технологию защиты документов от подделок. Сущность предлагаемой технологии заключается в том, что в конец документа добавляется аудиомаркер, хранящий в себе основную информацию защищаемого документа (финансовые аспекты, обязанности сторон, сроки и т.п.). Показано, что изменение текста в этом случае будет порождать необходимость изменения связанного с ним элемента защиты – аудиомаркера. Достоинством подхода является тот факт, что сделать это без автора или собственника документа невозможно, поскольку аудиомаркер хранит в себе биометрические данные указанного лица. Разработано приложение под мобильное устройство с операционной системой android, которое позволяет считывать, распознавать и воспроизводить аудиомаркеры. Проведенные испытания программного комплекса показали его высокую эффективность, в том числе, при распознавании аудиомаркеров в условиях различной освещенности. Рассмотрена возможность повышения степени защищенности подлинности документа за счет фиксации в речевой подписи текущего уровня психоэмоционального напряжения автора документа, что может быть использовано для выявления случаев его неадекватного состояния. Это позволяет выявлять случаи составления документов при применении физического или психического насилия.

Анастасия А. Баринова, Сергей В. Запечников

Методы и средства обеспечения конфиденциальности смарт-контрактов

Ключевые слова: смарт контракты, блокчейн, конфиденциальность

В настоящее время наблюдается тенденция качественного и количественного усложнения бизнес-процессов. Электронная форма обмена информацией становится преобладающим средством коммуникаций, а также одним из основных элементов обеспечивающей инфраструктуры для бизнес-процессов. Эти обстоятельства обеспечивают необходимость внедрения электронных смарт-контрактов («умных контрактов») – протоколов, которые

описывают набор условий и автоматически следят за их выполнением. Наиболее важным требованием к подобным технологиям является обеспечение конфиденциальности данных. В данной работе были исследованы наиболее известные платформы конфиденциальных электронных контрактов, выявлены методы обеспечения конфиденциальности, критерии их сравнения, проведено сравнение по выделенным критериям и сформулированы выводы о требованиях к платформам электронных контрактов, обеспечивающих конфиденциальность.

Филипп Б. Буртыка

Методы организации произвольных конфиденциальных запросов к криптографической облачной базе данных с помощью полностью гомоморфного шифрования

Ключевые слова: конфиденциальные запросы, криптографическая база данных, полностью гомоморфное шифрование, защищенные вычисления, секретные вычисления, недоверенная среда, криптографический протокол, конфиденциальное получение информации

Обработка запросов к зашифрованной базе данных без ее расшифрования в последнее время активно исследуется как специалистами по криптографии, так и исследователями в области БД. Такую обработку позволяют проводить различные виды т.н. обрабатываемого шифрования (ОШ), а также специальные архитектуры систем управления базами данных (СУБД), использующие эти виды шифрования. Самыми известными видами ОШ являются шифрование, сохраняющее порядок, гомоморфное шифрование, функциональное шифрование, поисковое шифрование и шифрование, сохраняющее свойства. На основе этих видов шифрований строятся различные СУБД, наиболее известными примерами которых могут служить CryptDB, Monomi, Arx и СУБД исследователей из Новосибирска. Эти СУБД строятся на применении комбинации различных видов ОШ, например шифрования сохраняющего порядок, гомоморфного шифрования и традиционных видов шифрования. Однако, такой подход может вызвать проблемы с конфиденциальностью. Наилучшим подходом с точки зрения защищенности было бы построение криптографической СУБД с использованием только гомоморфного шифрования. Препятствием к этому пока что является недостаточная эффективность самого гомоморфного шифрования а также неполное решение комплекса вопросов, связанных с необходимостью конфиденциального принятия решения в недоверенной среде. В данной работе предлагаются методы для решения этих вопросов, в частности для организации выполнения произвольного конфиденциального запроса к зашифрованной реляционной БД с помощью ПГШ. Также предлагается модель запроса к БД, разделяющая условие запроса на атомарные предикаты и связывающее условие. Один из предлагаемых методов направлен на сохранение конфиденциальности связывающего условия, другие на сохранение конфиденциальности атомарных предикатов. Параметры предложенных методов позволяют реализовать их с использованием уже существующих схем гомоморфного шифрования. Предлагаемые методы могут служить основой для построения систем защищенных криптографических облачных баз данных.

Виктор С. Горбатов, Алексей А. Мещеряков

Курс тренинга по безопасной разработке программного обеспечения

Ключевые слова: программное обеспечение, безопасность приложений, обучение, образование, информационная безопасность

Безопасность является одним из важнейших критериев качества разрабатываемого программного обеспечения. Для получения достаточного уровня защищенности разрабатываемых приложений компании внедряют процессы обеспечения безопасности в жизненных цикл разработки программного обеспечения. Так компании-разработчики

сталкиваются с дефицитом кадров, способных решать задачи как по проектированию и реализации ПО, так и по обеспечению их защиты. Данная статья предлагает описание курса тренинга по безопасной разработке программного обеспечения. Курс по безопасной разработке приложений предназначен для совместного обучения студентов, получающих образование как в области проектирования и реализации ПО, так и в области информационной безопасности.

Александр М. Коротин

Анализ угроз безопасности ответственной информации, передаваемой системой АЛС на базе радиоканала

Ключевые слова: автоматическая локомотивная сигнализация на базе радиоканала, АЛСР, безопасность АЛСР, модель угроз и нарушителя АЛСР, безопасность ответственной информации, обеспечение целостности ответственной информации

Переход к использованию радиоканала для передачи ответственной информации в системах автоматической локомотивной сигнализации (АЛС) усложняет обеспечение безопасности перевозочного процесса, так как в этом случае создаются условия для получения несанкционированного доступа к каналу связи из-за пределов контролируемой зоны. В связи с этим представляется целесообразным решение задачи по определению угроз безопасности ответственной информации, передаваемой системой автоматической локомотивной сигнализации на базе радиоканала (АЛСР), реализация которых способна нарушить безопасность перевозочного процесса. Решение поставленной задачи возможно путем анализа угроз безопасности ответственной информации, определения степени их влияния на перевозочный процесс, разработки модели нарушителя. В статье показано, что нарушение безопасности перевозочного процесса возможно в случае нарушения целостности ответственной информации. Используя методические документы ФСТЭК России, разработана модель нарушителя безопасности ответственной информации. Сделан вывод, что угрозы со стороны внутреннего нарушителя для традиционных систем автоматической локомотивной сигнализации и систем АЛСР совпадают. Сформирован перечень угроз безопасности ответственной информации, направленных на нарушение безопасности перевозочного процесса со стороны внешнего нарушителя. Полученные в статье результаты могут использоваться при исследовании безопасности систем АЛСР, анализе существующих решений по защите ответственной информации, передаваемой АЛСР, или разработке новых способов защиты.

Андрей Н. Кузнецов, Дмитрий А. Вышемирский

Об одном подходе к решению задачи токенизации при анализе больших массивов пользовательских паролей

Ключевые слова: пароли, утечки, токенизация, разбиение на слова

Проведен анализ алгоритма выделения в паролях пользователей слов на естественных языках (токенизации), предложенного ранее в работе R. Veras et al. [1] (RV-алгоритм). Выявлены основные недостатки данного подхода. Предложен новый алгоритм токенизации (RGramToken) на основе частотных словарей слов, биграмм и триграмм, позволяющий лучше учесть информацию о вероятности использования слов и словосочетаний в естественном языке. Подготовлены тестовые выборки фраз с учетом возможных искажений типа «вставка» между словами естественного языка. Проведен сравнительный анализ результатов работы на тестовых и реальных выборках алгоритма RGramToken с RV-алгоритмом. Показано более высокое качество токенизации предложенным в работе алгоритмом на искаженных фразах, а также устойчивость к качеству используемых словарей.

Владимир А. Минаев, Екатерина В. Вайц, Юлия В. Грачёва

Имитационные эксперименты с моделью информационно-психологических воздействий на массовое сознание

Ключевые слова: системно-динамическое моделирование, имитационное моделирование, имитационные эксперименты, информационно-психологические воздействия, массовое сознание

В статье рассматривается проблема исследования динамики информационно-психологических воздействий на массовое сознание, возможность их прогнозирования и управления, что является одним из важнейших аспектов обеспечения информационно-психологической безопасности общества и его граждан. Для исследования динамики информационно-психологических воздействий на массовое сознание в статье предложен метод системно-динамического моделирования, обоснованный и реализованный на моделях сложных социально-экономических явлений Дж. Форрестером в 1950-х годах. Применение данного метода к решению проблем обеспечения информационной безопасности исследуется различными зарубежными научными коллективами. Метод системно-динамического моделирования позволяет отобразить и исследовать множество факторов, критично влияющих на процессы информационно-психологических воздействий на массовое сознание. В предложенной авторами статьи модели учтены следующие факторы: вероятность «увлечения» идеями, заложенными в контент информационно-психологических воздействий, при межличностном контакте и в результате воздействия средств массовой информации, показатель массовости и регулярности средств массовой информации, пропагандирующих идею информационно-психологических воздействий, вероятность забывания идеи, заложенной в контент информационно-психологических воздействий, вероятность коммуникации на тему, заложенную в контент информационно-психологических воздействий, вероятность «увлечения» идеями информационно-психологических воздействий при одном коммуникационном контакте, среднее количество ознакомлений с одним сообщением в средствах массовой информации, вероятность «увлечения» новой идеей из контента информационно-психологических воздействий после ознакомления с сообщением в средствах массовой информации, количество контактов человека, увлеченного идеями информационно-психологических воздействий, в день. Математическим аппаратом, на котором базируются системно-динамические модели, являются системы дифференциальных уравнений. Авторами статьи построена имитационная системно-динамическая модель, позволяющая выявлять взаимосвязи между факторами, влияющими на динамику информационно-психологических воздействий на массовое сознание, и выполнена ее реализация в программной среде Anylogic. Данная среда является одной и самых перспективных сред имитационного моделирования. Приведены временные графики, отражающие динамику количества лиц, способных принять идею информационно-психологических воздействий, принявших идею информационно-психологических воздействий, а также графики, отражающие распределение во времени количества лиц, принявших идею информационно-психологических воздействий при определенных начальных условиях. Проведены имитационные эксперименты, связанные с варьированием параметров модели, в которых исследуется зависимость темпа увеличения количества лиц, принявших идею информационно-психологических воздействий, от частоты контактов лиц, разделяющих данные идеи.

Иван В. Нечта

Анализ устойчивости динамических водяных знаков

Ключевые слова: динамические водяные знаки, стеганография исполняемых файлов

В данной работе рассмотрена ранее известная схема внедрения динамических водяных знаков Radix-n, применяемая для борьбы с нелегальным использованием программного обеспечения. Согласно схеме водяной знак представляет собой динамическую структуру данных (граф), которую программа создает в памяти компьютера в процессе выполнения.

Топология такого графа определяет скрытую информацию, например, сведения об авторе. Эта информация может быть извлечена для установления авторства в рамках судебных разбирательств. В ходе данного исследования было установлено, что рассматриваемая схема, считавшейся ранее одной из наиболее надежных, обладает рядом особенностей, позволяющих злоумышленнику выявить этап построения водяного знака в программе и, соответственно, его исказить или удалить. Автором данной статьи показан недостаток схемы Radix-n, который заключается в том, что с помощью сведений, полученных от перехватчика Api-функций выделения динамических блоков памяти, можно выявить динамические структуры данных программы, среди которых присутствует водяной знак. Анализ содержимого оперативной памяти позволяет обнаружить указатели на динамически созданные объекты (массивы, переменные, экземпляры классов и т.д.) программы. Различные динамические объекты в памяти, связанные между собой указателями образуют структуры данных программы (списки, стеки, деревья и иные графы), в том числе и водяной знак. В ходе проведенных экспериментов установлено, что в подавляющем числе случаев количество структур данных в программах невелико, что повышает вероятность реализации успешной атаки злоумышленником. В настоящей работе также предложен алгоритм поиска компонент связности графа с линейной трудоемкостью, в случае, когда количество вершин порядка 106. На основе полученных в эксперименте результатов была предложена новая схема внедрения водяных знаков, являющаяся устойчивой к разработанной атаке. В предлагаемой схеме используется граф для хранения водяного знака, где связи между узлами реализованы с помощью уникальных сигнатур. Предложено использовать шифрование содержимого узлов графа (кроме сигнатур) для того, чтобы злоумышленник не смог выявить наличие связи между узлами.

Алина В. Трепачева

О соотношениях между атаками на симметричные шифры, гомоморфные над кольцом вычетов

Ключевые слова: гомоморфное шифрование над кольцом вычетов, полностью гомоморфное шифрование, задача RSA, задача факторизации чисел, атака только по шифртекстам, атака по известным открытым текстам

Работа посвящена изучению криптостойкости симметричных гомоморфных шифров над кольцами вычетов. Основной задачей является выяснить, возможно ли установить эквивалентность между атакой только по шифртекстам (АТШ) и атакой по известным открытым текстам (АИОТ) для таких шифров. Для этого вводится понятие сводимости между атаками и дается достаточное условие сводимости от АТШ к АИОТ. Основная идея заключается в том, что для доказательства сводимости от АТШ к АИОТ необходимо найти функцию над кольцом вычетов, которая является эффективно вычислимой и имеет небольшой размер образа по сравнению с размером всего кольца вычетов. Исследование наличия сводимости интересно тем, что оно может позволить лучше понять уровень криптостойкости существующих симметричных гомоморфных криптосистем (ГК). Поскольку для большинства из них в литературе уже установлена уязвимость к АИОТ, то доказательство существования сводимости может показать, что эти криптосистемы не стойки даже к АТШ, а, следовательно, совсем не стойки и непригодны для применения. Приводится пример сводимости АТШ к АИОТ для случая, когда кольцо вычетов является простым полем. На основе этого примера описывается эффективная АТШ на одну симметричную ГК в случае кольца вычетов небольшого размера. Также отдельно рассматривается случай, когда кольцо вычетов для ГК строится по труднофакторизуемому модулю n. Для таких n на данный момент не известен эффективный алгоритм построения эффективно вычислимых функций с «небольшим» образом. Ввиду этого дальнейшая работа по криптоанализу существующих симметричных ГК будет направлена на изучение свойств функций над кольцами вычетов по труднофакторизуемым модулям.

SUMMARY

Alexandr M. Alyushin, Sergey V. Dvoryankin

The use of speech technology to protect the document turnover

Keywords: spectrogram, audiomarker, recognition, documentation, graphic image

The wide current paper documents implementation in practice workflows are shown. The basic aspects of document protection related to the protection of their content and legal components are underlined. For contextual component assigned semantic information aspect of the document is considered. For legal component attributed facts and conditions for the creation, approval, negotiation of the document to specific persons is viewed. The documents protection problem importance is shown in connection with possible terrorist threats. The importance of such factor as the time of fraud detection towards the efficiency of documents protection is shown. The fraud detection time requirements for documents of different nature - financial, legal, management is analyzed. The documents used for the operational management of dangerous objects is point out as the most sensitive to the falsification. It is shown that their deliberate falsification can lead to accidents and technogenic catastrophes and human casualties. A comparative analysis of currently used protecting documents methods are presented. Biometric and non-biometric methods of documents protection are point out.The analysis of their short comings are given. The conclusion about the prospects of document protection on the basis of the voice signature technology are done. The basic steps of voice information processing in the implementation of this technology are analyzed. The software that implements a documents counterfeiting new protection technology is proposed. The technology is based on the audiomarkers usage at the end of the document, which contains a general information about it. The technology is applicable to the wide range of documents such as financial and valuable papers, contracts, etc. One of the most important advantages of this technology is that any changes in the document can not be done without the author of the document because audiomarker keeps the biometric data of the person. The developed software is oriented on the mobile applications on the basis of android operating system. The software has a friendly interface, which allows one to create, correct and check audiomarkers. The performed software tests shows its high efficiency, including the possibility of audiomarkers recognition in various lighting conditions. The possibility of document protection level improvement by logging in voice signature the current level of mental and emotional stress of the document author, which can be used to detect its inadequate state is shown. This makes it possible to identify cases of document preparation under the application of physical or mental violence.

Anastasia A. Barinova, Sergey V. Zapechnikov

On the techniques and tools for privacy-preserving smart contracts

Keywords: smart contracts, privacy, blockchain

Currently, business processes become more and more complicated. Data used in these processes circulates mainly through the digital communications. Due to these conditions some kind of electronic contracts for business deals becomes necessary. Smart contracts should describe a set of conditions, implemented through some events in the real world and digital systems. The most important requirement for this technology is privacy ensuring. In this work we have explored existing projects of privacy-preserving smart contracts, defined comparison criteria, compared projects and made a conclusion about options required for smart contract frameworks.

Filipp B. Burtyka

The Techniques for Arbitrary Secure Quering to Encrypted Cloud Database Using Fully Homomorphic Encryption

Keywords: Private Queries, Encrypted Database, Fully Homomorphic Encryption, Secure Computations, Secret Computations, Untrusted Environment, Cryptographic Protocol, Private Information Retrieval, Querying Encrypted Data

The processing of queries to an encrypted database without its decryption has been actively investigated recently by both cryptographers and database researchers. Such a work is allowed by various types of so-called Processable encryption (PE), as well as special architectures of database management systems (DBMS) which use these types of encryption. The most known types of PEs are order-preserving encryption, homomorphic encryption, functional encryption, searchable encryption, and property-preserving encryption. Based on these types of encryption, various DBMSs are built, the most famous of which are CryptDB, Monomi, Arx and DBMS by researchers from Novosibirsk. These DBMSs are built on the basis of various types of PEs, for example order-preserving encryption, homomorphic encryption and traditional block encryption. However, this approach can cause privacy problems. The best approach from the security viewpoint is to build a cryptographic database using only homomorphic encryption. An obstacle to this is insufficient efficiency of the existing homomorphic encryption schemes and incomplete solution of a set of issues related to ensuring the confidentiality of decisions making in an untrusted environment. In this paper, we propose the techniques for solving these problems, in particular for organization of execution arbitrary secure query to the encrypted relational database using fully homomorphic encryption. Also we propose a model of query condition that splits query into atomic predicates and linking condition. One of roposed technique is aimed at ensuring the security of linking condition of queries, others keep security of atomic predicates. The parameters of the proposed techniques make it possible to implement them using the already existing homomorphic encryption schemes. The proposed techniques can be a basis for building secure cryptographic cloud databases.

Viktor S. Gorbatov, Alexey A. Meshcheriakov

Secure software development training course

Keywords: information security, education, secure software development, application security, training

Information security is one of the most important criteria for the quality of developed software. To obtain a sufficient level of application security companies implement security process into software development life cycle. At this stage software companies encounter with deficit employees who able to solve problems of software design, implementation and application security. This article provides a description of the secure software development training course. Training course of application security is designed for co-education students of different IT-specializations.

Alexander M. Korotin

Security threat analysis of critical information transmitted by the ALS system based on the radio channel

Keywords: automatic locomotive signaling based on radio channel, ALSR, security of ALSR, threats model and violator model of ALSR, security of critical information, ensuring of critical information integrity

The transition to the use of a radio channel for the transfer of critical information in automatic locomotive signaling (ALS) systems complicates the security of the transportation process, because in this case conditions are created for obtaining unauthorized access to the communication channel from outside the controlled area. That's why it seems advisable to solve the problem of

identifying security threats of critical information transmitted by the automatic locomotive signaling system based on the radio channel (ALSR), the implementation of which is likely to disrupt the safety of the transportation process. The solution of this task is possible by analyzing security threats of critical information, determining the degree of their influence on the transportation process and developing violator model. The article shows that a security breach of the transportation process is possible in case of violation of the responsible information integrity. A model of violator of critical information security was developed using methodical documents of FSTEC of Russia. It was concluded that the threats from the internal violator for traditional automatic locomotive signaling systems and ALSR systems coincide. A list of security threats of sensitive information was developed, directed to violating the security of the transportation process by an external violator. The results obtained in the article can be used to study the safety of ALSR systems, to analyze existing solutions for protecting critical information transmitted by ALSR, or to develop new methods of protection.

Andrey N. Kuznetsov, Dmitry A. Vyshemirsky

One approach to solving tokenization problem for analysis of large-scale collections of user-defined passwords

Keywords: passwords, leaks, tokenization, words isolation

This paper performs an analysis of the algorithm of password tokenization introduced by R. Veras et al. [1]. We show main limitations of this approach and propose a new tokenization algorithm - RGramToken, based on frequency dictionaries of English words, bigrams and trigrams. Our approach allows better utilization of information about probabilitiy distribution of words and word combinations in a natural language. The results of comparison analysis of these two algorithms on specially prepared tests with warped phrases demonstrate higher efficiency of RGramToken and its robustness on low quality dictionaries.

Vladimir A. Minaev, Ekaterina V. Vaitc, Yulia V. Gracheva

Simulation experiments with the model of information-psychological influences on mass consciousness

Keywords: system-dynamic modeling, imitating simulation, simulation experiments, information and psychological influence, mass consciousness

The article deals with the problem of researching the dynamics of information and psychological influences on mass consciousness, the possibility of their forecasting and management, which is one of the most important aspects of ensuring the information and psychological security of society and its citizens. To research the dynamics of information and psychological influence on mass consciousness, the article suggests a method of system-dynamic modeling, grounded and implemented on models of complex socio-economic phenomena by J. Forrester in the 1950s. The application of this method to solving the problems of information security is investigated by various foreign scientific teams. The method of system-dynamic modeling allows to display and investigate many factors that critically affect the processes of information and psychological influences on mass consciousness. The following factors are taken into account in the model proposed by the authors of the model: the probability of "enthusiasm" of ideas embedded in the content of information and psychological influences, with interpersonal contact and as a result of the influence of the mass media, the massiveness and regularity of the mass media propagandizing the idea of information and psychological influences, the probability of forgetting the idea, embedded in the content of information and psychological influences, the probability of communication on the topic, embedded in the content of information and psychological influences, the probability of "enthusiasm" ideas of information and psychological influences in one communication contact, the average number of acquaintances with one message in the media, the probability of "enthusiasm" of a new idea from the content of information and psychological influences after reading the information in the media, the number of contacts a person, keen on ideas of information and psychological influences, per day. The mathematical apparatus on which

system-dynamic models are based are systems of differential equations. The authors of the article constructed an imitation model that makes it possible to identify the interrelationships between factors that effect the dynamics of information and psychological influences on mass consciousness, and its implementation has been carried out in the Anylogic software environment. This environment is one of the most promising environments for simulation. Temporal graphs reflecting the dynamics of the number of persons capable of adopting the idea of information and psychological influence, the number of persons who adopted the idea of information and psychological influences, and graphs reflecting the distribution in time of the number of persons who adopted the idea of information and psychological influences under certain initial conditions are given. Simulation experiments related to variations of the model parameters in which the dependence of increasing rate of person's number who adopted the idea of information and psychological effects on the frequency of person's contacts sharing these ideas is studied.

Ivan V. Nechta

Robustness analysis of dynamic watermarks

Keywords: dynamic watermarking, steganography of executables

In this paper we consider previously known scheme of dynamic watermarks embedding (Radix-n) that is used for preventing illegal use of software. According to the scheme a watermark is dynamic linked data structure (graph), which is created in memory during program execution. Hidden data, such as information about author, can be represented in a different type of graph structure. This data can be extracted and demonstrated in judicial proceedings. This paper declared that the above mentioned scheme was previously one of the most reliable, has a number of features that allows an attacker to detect a stage of watermark construction in the program, and therefore it can be corrupted or deleted. The author of this article shows the weakness of Radix-N scheme, which consists in the fact that we can reveal dynamic data structures of a program by using information received from some API-functions hooker which catches function calls of dynamic memory allocation. One of these data structures is the watermark. Pointers on dynamically created objects (arrays, variables, class items, etc.) of a program can be detected by content analysis of computer`s RAM. Different dynamic objects in memory interconnected by pointers form dynamic data structures of a program such as lists, stacks, trees and other graphs (including the watermark). Our experiment shows that in the vast majority of cases the amount of data structure in programs is small, which increases probability of a successful attack. Also we present an algorithm for finding connected components of a graph with linear time-consuming in cases where the number of nodes is about 106. On the basis of the experimental findings the new watermarking scheme has been presented, which is resistant to the proposed attack. It is offered to use different graph structure representation of a watermark, where edges are implemented using unique signatures. Our scheme uses content encrypting of graph nodes (except signature), so an attacker is not able to reveal links between nodes and find the watermark.

Alina V. Trepacheva

On the relations between the attacks on symmetric homomorphic encryption over the residue ring

Keywords: ring-homomorphic encryption, fully homomorphic encryption, RSA problem, number factorization problem, ciphertexts only attack, known plaintext attack

The paper considers the security of symmetric homomorphic cryptosystems (HC) over the residue ring. The main task is to establish an equivalence between ciphertexts only attack (COA) and known plaintexts attack (KPA) for HC. The notion of reducibility between attacks and sufficient condition of reducibility from COA to KPA are given for this purpose. The main idea is: to prove reducibility from COA to KPA we need to find a function over residue ring being efficiently computable and having a small image size comparing with the size of residue ring. The study of reducibility existence is important since it allows to understand better the security level of

symmetric HC proposed in literature. A vulnerability against KPA has been already found for the majority of these HC. Thus the reducibility presence can demonstrate that cryptosystems under the study are not secure even against COA, and therefore they are totally insecure and shouldn't be used in practice. We give an example of reducibility from COA to KPA for residue ring being a simple field. Based on this example we show an efficient COA on one symmetric HC for small field. Also we separately consider the case of residue ring composed using number n being hard-to-factor. For such n an efficient algorithm to construct an efficiently computable function with small image is unknown so far. So further work related to cryptanalysis of existing symmetric HC will be directed into study of functions properties over residue rings modulo numbers hard for factorization.

ПРАВИЛА ДЛЯ АВТОРОВ

Рукописи, предоставляемые в редакцию, должны соответствовать следующим требованиям:

- тема статьи должна быть актуальной, иметь научное или практическое значение и публиковаться авторами впервые;
- рукопись должна быть оформлена только в формате *.doc, полоса А4, кегль 12, шрифт Times New Roman, интервал полуторный;
- в начале статьи идут сведения о статье **на английском языке**: И.О. Фамилия авторов (по центру, строчными буквами); далее сведения об авторах - должность, ученая степень, ученое звание, место работы, контактный телефон, адрес электронной почты и личный идентификатор ORCID (по центру, строчными буквами, курсив); затем название статьи (по центру, строчными буквами, полужирно с подчеркиванием); ключевые слова (не более шести, по ширине, курсив); аннотация (8-12 строк, по ширине, строчными буквами);
- далее идут сведения о статье **на русском языке**: И. О. Фамилия авторов (по центру, строчными буквами); далее сведения об авторах - должность, ученая степень, ученое звание, место работы, контактный телефон, адрес электронной почты и личный идентификатор ORCID (по центру, строчными буквами, курсив); затем название статьи (по центру, строчными буквами, полужирно с подчеркиванием); ключевые слова (не более шести, по ширине, курсив); аннотация (8-12 строк, по ширине, строчными буквами);
- затем идет текст статьи на русском или английском языке, кегль 12, интервал полуторный, рекомендуемый общий объем статьи не должен превышать 10 страниц, включая таблицы, иллюстрации;
- в конце статьи приводится СПИСОК ЛИТЕРАТУРЫ, в котором указан библиографический список источников литературы, оформленный в соответствии с действующими стандартами (как правило, не менее 15 наименований);
- после списка литературы идет REFERENCES, в котором эти библиографические источники должны быть написаны латиницей (т. е. латинскими буквами).

Условия опубликования статьи:

- статья может быть выслана по электронной почте или представлена в редакцию на бумажном (одном экземпляре) и электронном носителях (кроме дискет);
- редакционная коллегия журнала следует этическим нормам, принятым в международном научном сообществе, опираясь на рекомендации Комитета по этике научных публикаций, не противоречащим нормам российского законодательства в областях регулирования деятельности средств массовой информации и авторского права;
- статьи, не соответствующие установленным требованиям представления и оформления, не рассматриваются и не публикуются;
- в одном номере журнала публикуется, как правило, только одна статья автора, в том числе с соавторами;
- авторы должны предоставлять только оригинальные работы, при использовании текстовой или графической информации, полученной из работ других лиц, необходимы ссылки на соответствующие публикации или письменное разрешение автора;
- решение о публикации рукописи принимается редакционной коллегией на основании результата рецензирования и экспертной оценки квалифицированными специалистами в области ИБ;
- в случае приема рукописи к публикации автор должен оперативно давать ответы на

вопросы редакции, связанные с замечаниями по статье;
- в случае отказа в публикации редакционная коллегия должна предоставить автору копию рецензии и обоснование отказа публикации;
- подача статьи в более чем один журнал одновременно расценивается как неэтичное поведение и является неприемлемой;
- статьи публикуются бесплатно.

Правила оформления текстов для публикации

1. Статьи необходимо подавать в электронном виде (*.doc или *.rtf) с распечаткой (или файлом в формате *.pdf) - во избежание неточностей прочтения формул.
2. Картинки, графики, фотографии и другие виды иллюстраций, по возможности, следует предоставлять не только включенными в текст, но и отдельными файлами в исходном формате (не интегрированными в документ Word).
3. Сокращения и аббревиатуры, которых нет в списке сокращений, необходимо раскрывать (в скобках или в сноске).
4. Давая в тексте статьи ссылки на формулы, выражения или ограничения, пожалуйста, убедитесь в том, что соответствующие объекты в статье есть и пронумерованы.
5. Ссылки на литературу следует давать в тексте в квадратных скобках, в случае цитирования - с указанием страниц.
6. При оформлении списка литературы желательно обращать внимание на наличие выходных данных работ и избегать повторных указаний одной и той же работы под разными номерами.
7. Ссылки на законы, нормативные акты, конфессии и прочее желательно указывать по установленной форме: Закон РФ « » от х месяца хххх г. № ___. Ст.___.
8. Иноязычные слова, термины и фамилии, написание которых допускает варианты, просьба писать в пределах одной статьи одинаково.

Заранее спасибо,
редакционная коллегия

AUTHOR GUIDELINES

The articles submitted to the editors must meet the following requirements:

– the topic of the article should be relevant, have scientific or practical significance and be published by the authors for the first time;
– the manuscript should be formatted only in * .doc or pdf format, A4 strip, size 12, TimesNewRoman font, one-and-a-half interval;
– in the beginning of the article there are information about the article in English: I.O. Name of authors (centered, lower case); Further information about authors – position, academic degree, academic title, place of work, contact phone number, e-mail address and personal identifier ORCID (centered, lowercase, italics); Then the title of the article (centered, lowercase, bold with underline); Keywords (no more than six, in width, italics); Annotation (8–12 lines, width, lower case);
– Further information on the article is in Russian: I.O. The authors' surname (for jubilus, lower case letters); Further information about authors – position, academic degree, academic title, place of work, contact phone number, e-mail address and personal identifier ORCID (centered, lowercase, italics); Then the title of the article (centered, lowercase, bold with underline); Keywords (no more than six, in width, italics); Annotation (8–12 lines, width, lower case);
– then the text of the article is in Russian or English, size 12, interval one and a half, the recommended total volume of the article should not exceed 10 pages, including tables, illustrations;
– at the end of the article the LIST OF LITERATURE is given, in which the bibliographic list of sources of literature is indicated, drawn up in accordance with the current standards (as a rule, not less than 15 titles);
– after the list of literature is REFERENCES, in which these bibliographic sources should be written in Latin (ie Latin letters).

Terms of publication of the article:

– the article should be sent by e-mail;
– The editorial board of the journal follows the ethical standards adopted in the international scientific community, relying on the recommendations of the Ethics Committee of scientific publications that do not contradict the norms of Russian legislation in the field of regulation of the activities of the media and copyright;
– articles that do not meet the requirements for presentation and processing are not considered or published;
– in one issue of the journal, as a rule, only one author's article is published, including coauthors;
– authors should provide only original works, if text or graphic information obtained from other persons is used, references to the relevant publications or the author's written permission are necessary;
– the decision to publish the manuscript is made by the editorial board on the basis of the result of peer review and expert evaluation by qualified specialists in the field of information security;
– in the case of receipt of the manuscript for publication, the author must promptly give answers to editorial questions related to comments on the article;
– in case of refusal to publish, the editorial board should provide the author with a copy of the review and justification for refusing the publication;
– Submitting an article to more than one journal is simultaneously regarded as unethical behavior and is unacceptable;
– articles are published for free.

Rules for publication of texts

1. Articles must be submitted electronically (* .doc or * .rtf) with a printout (or a file in * .pdf format) - to avoid inaccuracies in reading the formulas.
2. Pictures, graphics, photographs and other types of illustrations should, if possible, not only be included in the text, but also separate files in the original format (not integrated into the Word document).
3. Abbreviations and abbreviations, which are not on the list of abbreviations, should be disclosed (in parentheses or in a footnote).
4. By providing links to formulas, expressions or restrictions in the text of the article, please make sure that the relevant objects in the article are numbered and numbered.
5. References to the literature should be given in the text in square brackets, in the case of citations, with pages.
6. When preparing a list of literature, it is desirable to pay attention to the availability of output data of works and to avoid repeated instructions of the same work under different numbers.
7. References to laws, regulations, confessions and so on should be indicated in the prescribed form: the Law of the Russian Federation «__» of x month xxxx, No.___. Art.___.
8. Foreign words, terms and surnames, the spelling of which allows variants, please write within the same article the same way.

Submission Preparation Checklist

As part of the submission process, authors are required to check off their submission's compliance with all of the following items, and submissions may be returned to authors that do not adhere to these guidelines.

1. This article has not been previously published, and not submitted for review and publication in another journal (or a corresponding explanation if otherwise in the Comments to the editor).
2. File with the articles submitted in the one jf the following document format OpenOffice, Microsoft Word, RTF, or WordPerfect.
3. The full web address (URL) for links are given where it is possible.
4. The text is single-spaced; uses a font size of 12 points; to highlight use italics, not underlining (except for URL addresses); all illustrations, graphs and tables located in the appropriate places in the text, not at the end of the document.
5. The text complies with the stylistic and bibliographic the requirements described in the Guide for authors, on the «About the journal» page.
6. If you are submitting an article in a peer reviewed section of the journal then the document meets the requirements to ensure blind peer review.

Privacy Statement

The names and email addresses entered in this journal site page will be used exclusively for the purposes specified by this journal and will not be used for any other purposes or will not be given over to another individuals and organizations.

СПИСОК СОКРАЩЕНИЙ, ПРИНЯТЫХ В ЖУРНАЛЕ

АБИ – администратор безопасности информации
АнД – аналоговый документ
АРМ АБИ – автоматизированное рабочее место администратора безопасности информации
АС – автоматизированная система
БД – база данных
БИС – большая интегральная схема
ИБ – информационная безопасность
ИКТ – информационно-коммуникационные технологии
ИП – информационные продукты
ИПС – изолированная программная среда
ИР – информационные ресурсы
КСЗ – комплекс средств защиты
КТЭ – компьютерно-техническая экспертиза
ЛВС – локальная вычислительная сеть
МЭ – межсетевой экран
НД – нормативный документ
НСД – несанкционированный доступ
ОИ – объект информатизации
ОС – операционная система
ПАК – программно-аппаратный комплекс
ПО – программное обеспечение
ПРД – правила разграничения доступа
ПСКЗИ – персональное средство криптографической защиты информации
РД – руководящий документ
РКБ – резидентный компонент безопасности
РПВ – разрушающее программное воздействие
СВТ – средство вычислительной техники
СЗИ – средство защиты информации
СЗИ НСД – средство защиты информации от несанкционированного доступа
СКЗИ – система криптографической защиты информации
СРД – система разграничения доступа
СУБД – система управления базами данных
ЭлД – электронный документ
ЭЦП – электронная цифровая подпись

Адрес редакции: Каширское шоссе, 31, Москва, 115409, Россия
Тел.: +7 (495) 788 5699, тоновый режим 9216 или 9087.
Факс: +7 (499) 324-86-00.
Editorial address: Kashirskoe shosse, 31, Moscow, 115409, Russia
Tel. +7 (495) 788 5699, tone mode set 9216 or 9087.
Fax: +7 (499) 324-86-00.
E-mail: BIT@mephi.ru
https://bit.mephi.ru

Периодичность выхода – 4 раза в год / Periodicity – 4 times a year

Подписка на журнал
производится на почтовых отделениях связи
по каталогу «Пресса России»

Подписной индекс 29226

Цена в продаже свободная / Price selling free

Технический редактор Т.В. Волвенкова
Корректор Е.Г. Станкевич
Верстка Г.А. Бобровой

Подписано в печать 11.05.2017. Формат 60x84 1/8.
Печ.л. 13,25. Уч.-изд.л. 13,25. Тираж 500 экз. Изд. № 002-3.

Национальный исследовательский ядерный университет «МИФИ».
Каширское шоссе, 31, Москва, 115409, Россия.
National Research Nuclear University MEPHI.
Kashirskoe shosse, 31, Moscow, 115409, Russia

Типография ООО «Клуб печати».
127018, Москва, Марьиной Рощи 3-й проезд, 40