УДК 004.032.26+519.237.5
ББК 22.18
Ч 48

Черкасов А.Н.
Кубанский государственный технологический университет, Краснодар, Россия, cherk@mail.ru
Туркин Е.А.
Кубанский государственный технологический университет, Краснодар, Россия, wolfutur@mail.ru
Чемулова А.В.
Кубанский государственный технологический университет, Краснодар, Россия, chemulovaa@yandex.ru

Разработка вербальной модели кибердевианта при исследовании онлайн-сообществ на основе мультиномиальной логистической регрессии

(Рецензирована)

Аннотация. *Теоретический анализ актуальности существующих работ по изучению человеческого фактора в сфере информационной безопасности позволил разработать модель мультиномиальной логистической регрессии, позволяющей выполнять задачи классификации пользователей на злоумышленников различного ранга на основе текстовой аналитики Интернет-форума. С учетом особенностей полученных данных разработан алгоритм автоматического извлечения ключевых слов из сообщений с помощью алгоритма деревьев решений. В процессе обучения точность данного алгоритма достигла наиболее высокого по сравнению и другими алгоритмами (SVM, k-NN) значения – 78,74%. Произведенное моделирование позволило создать анкету для проведения психолингвистического эксперимента по определению уровня склонности человека к совершению киберпреступления. Ключевой аспект значимости проведенной работы заключается в апробации применения нетрадиционного для современной науки исследовательского подхода, который состоит в интеграции социологии, психолингвистики и статистического моделирования.*

Ключевые слова: *информационная безопасность, человеческий фактор, мультиномиальная логистическая регрессия, Интернет-форум, кибердевиант, классификация*

Cherkasov A.N.
Kuban State University of Technology, Krasnodar, Russia, cherk@mail.ru
Turkin E.A.
Kuban State University of Technology, Krasnodar, Russia, wolfutur@mail.ru
Chemulova A.V.
Kuban State University of Technology, Krasnodar, Russia, chemulovaa@yandex.ru

Development of a verbal cyber deviant model in the study of online communities based on multinomial logistic regression

Abstract. *A theoretical analysis of the relevance and existing works on the study of the human factor in the field of information security has been carried out. On its basis a model of multinomial logistic regression has been developed, which performs tasks of classification of users into intruders of different rank based on text analysis of the Internet forum. Taking into account the peculiarities of the obtained data, the algorithm of automatic extraction of keywords from messages using the decision tree algorithm has been developed. During the training process, the accuracy of this algorithm reached the highest value in comparison with other algorithms (SVM, k-NN) – 78.74%. The modelling carried out allowed creating a questionnaire for carrying out a psycholinguistic experiment to determine the level of tendency of a person to commit cybercrime. A key aspect of the significance of the work carried out was to test the adoption of a research approach that was unconventional for modern science, which consisted in the integration of sociology, psycholinguistics and statistical modeling.*

Keywords: *information security, human factor, multinomial logistic regression, Internet forum, cyber-deviant, classification*

Введение

В условиях автоматизации бизнес-процессов, развития новых технологий, а также построения новых каналов передачи информации риски информационной безопасности становятся ключевой проблемой для организаций. В частности, анализируя последние данные статистики, стоит отметить, что, по оценкам *Positive Technologies*, число целенаправленных атак уже превысило количество массовых и продолжает расти (59% во II квартале 2019 г.) [1]. При этом более половины всех киберпреступлений совершаются с целью кражи информации, а персональные данные – основной тип украденной информации в атаках на юридические лица (29%).

По данным опроса Egres, который был проведен среди руководящего персонала отделов кибербезопасности 500 IT-компаний, все чаще встречаются случаи возникновения угроз от инсайдеров компании. Половина случаев взлома кибербезопасности произошла по вине сотрудников. Обращаясь к проблеме нахождения путей решения данных происшествий, проведённое компанией Bitdefender исследование показало, что 38% руководителей компаний в области ИТ согласны с тем, что лучший способ защититься от сложных атак – это, прежде всего, эффективная работа с персоналом. Все вышеперечисленное предопределило необходимость моделирования новых путей анализа и прогнозирования роли человеческого фактора в сфере информационной безопасности.

В настоящее время исследования, позволяющие понять коренные особенности поведения тех людей, которые намерены совершить противоправные действия, являются не только междисциплинарной, но и малоизученной областью, особенно в российской практике. На данный момент такие крупные компании, как *SearchInform*, *Infowatch*, *Positive Technologies*, *ИнфоТеКС*, уже разрабатывают платформы, системы поддержки принятия решений для мониторинга поведенческих моделей сотрудников компании, также используя как новые методы текстовой аналитики, например, в рамках DLP-систем, так и в рамках ситуационных центров крупных корпораций [2].

Но основным недостатком подобных систем является отсутствие долговременного прогнозирования в области выявления потенциальных нарушителей информационной безопасности.

В свою очередь, в масштабах мирового пространства исследования информационной безопасности уже активно интегрируются в сферу социальной психологии, психолингвистики, а также социологии. Работы по использованию data- и text-mining, в частности анализ онлайн-сообществ, форумов, дают возможность построить социально-психологический портрет злоумышленника, например, хакера. Злонамеренные пользователи активно используют Интернет-форумы [3, 4]. Например, в исследовании [4] темы изучения сообщества злоумышленников делятся на три основных направления:

– качественный анализ для понимания и описания деятельности злоумышленника [5, 6];

– анализ сообществ теневой экономики [3, 6, 7];

– анализ и идентификация ключевых участников сообществ злоумышленников [8].

Одним из важных моментов, отмеченных в уже существующих исследованиях, является то, что часто участники Интернет-сообщества накапливают необходимые ресурсы, навыки и активы, чтобы сформировать группы для выполнения поставленной задачи-запроса [9, 10]. В группах разработчиков программного обеспечения, таких как форумы разработчиков iOS и Android, а также в сообществах злонамеренных пользователей относительно неопытные участники стремятся получить помощь от профессионалов и экспертов [3]. Существующие исследования анализируют поведение экспертов в области обхода информационных систем, но игнорируют изучение возможностей форума и моделей, основанных на текстовой аналитике [3, 11–13].

Профессионалы могут коммуницировать с другими пользователями социальных сетей через формирование однодольных и двудольных связей в этих форумах [14]. В рамках двудольных графов сами пользователи с другими пользователями не связаны. Связываются злоумышленники с группами на базе матрицы инциденций (связь между вершиной и ребром графа). Противоположностью двудольному графу является однодольный: в данный ситуации пользователи буду связаны в сети непосредственно с пользователями, а не с группами. Формируя подобные однодольные и двудольные связи, профессиональные злоумышленники пытаются найти лучших кандидатов на форумах.

В доступной литературе ([3, 4, 11]) функции форумов рассматривались только как основные факторы репутации и компетентности злоумышленников. Одним из главных недостатков существующих разработок является отсутствие разработанного инструмента психолингвистической оценки сотрудника на модель потенциального нарушителя с последующей классификацией и системой рекомендаций по работе с выявленным типом.

Таким образом:

– во-первых, лишь немногие исследования пытались классифицировать сообщества злоумышленников на основе их ролей и обязанностей, отдельно для каждой группы пользователей;

– во-вторых, не существует универсального вербального профиля злоумышленника, через который можно было бы анализировать его поведение; зачастую во внимание берется лишь поведенческая сторона профиля без лексических аспектов;

– в-третьих, проводится множество аналитических операций без проработанной схемы практического применения результатов, реального предотвращения киберпреступности.

Методы и материалы исследования

В связи с этим целью исследования стало конструирование анкеты для проведения лингвистического ассоциативного эксперимента. Анкета разработана на базе классификации вербальных профилей злоумышленников, которые формируются по результатам процедур интеллектуального анализа текста онлайн-форумов. Таким образом, исследование носит комплексный характер определения потенциального злоумышленника. Концептуальное представление схемы интеллектуального анализа текста онлайн-форумов представлено на рисунке 1.

Рис. 1. Концептуальное представление схемы интеллектуального анализа текста онлайн-форумов

Для разработки дорожной карты исследования следует определить само понятие «нарушитель информационной безопасности». Согласно ГОСТ Р 50922-2006, пункт 2.6.4. – «нарушитель информационной безопасности организации; нарушитель ИБ организации: физическое лицо или логический объект, случайно или преднамеренно совершивший действие, следствием которого является нарушение информационной безопасности организации».

Фрагментарно базируясь на общепринятом варианте определения нарушителя информационной безопасности и исходя из специфики поставленных целей, было сформулировано новое понятие – «кибердевианта», которое легло в основу разработанной методической карты по выявлению потенциального нарушителя в организации. Кибердевиант – человек с отклоняющейся в отрицательную сторону моделью поведения (с намерениями обойти систему и совершить противоправное действие), имеющий определенные лидерские качества, а также знания в области информационной безопасности.

Для конструирования и оценки вербального профиля кибердевианта разработана модель, которая учитывает ключевые слова и индекс их средней тональности. Главная цель заключается в анализе и выявлении факторов, значимых в рамках определения вербальных характеристик злоумышленников.

Каждое опубликованное злоумышленником сообщение связано с темой обсуждения информационной безопасности. Посредством textmining создаётся лексический портрет злоумышленника. Затем извлекается содержание настроений сообщения форума и применяется для создания классификации злонамеренных пользователей на основе их поведенческих ролей посредством модели мультиномиальной логистической регрессии. Для того чтобы сделать модель более надежной, в зависимость добавляются не только ключевые слова и тональность, но и более детальные параметры (длина сообщения и т.д.), способные исключить погрешности и ошибки.

Существующие исследования подтверждают, что количество символов, затраченных на доставку сообщения, сильно влияет на контент, создаваемый пользователем [11, 14]. Часто средняя длина сообщения для каждого сообщения используется как ковариата (количественная независимая переменная-предиктор). В ходе первичного анализа было обнаружено, что относительно длинные сообщения имеют большую познавательную ценность и гораздо более важны для более широкой аудитории сообщества. Количество слов и длина сообщений значительно увеличиваются на всех типах информационных уровней с увеличением глубины сообщений на обучающей онлайн-платформе. Исследования также подтверждают, что пользователи обычных социальных сетей, таких как Facebook, получают больше просмотров и ответов на более длинные сообщения. Аналогичные результаты наблюдаются и среди корпоративных социальных сетей (ESN)-s [15]. В условиях предприятия длина сообщения увеличивается для менеджеров и значительно уменьшается для других сотрудников при использовании электронной почты в качестве средства коммуникации. Такая ролевая демаркация также может быть экстраполирована на форумы злоумышленников с учетом уровня компетентности пользователя. Таким образом, определяется средняя длина сообщения в рамках среднего содержания символов для каждого пользователя. Поэтому в рамках данного исследования предполагается следующее:

– средняя длина сообщения определяет опытность злоумышленника;

– количество публикуемых сообщений определяет компетентность злоумышленника;

– суммарное количество тематических блоков, в которых участвует пользователь, определяет компетентность злоумышленника.

Более того важно содержание ключевых слов в сообщении, которое потенциально может являться значимым фактором, определяющим компетентность пользователя.

Индекс настроений определяет общее отношение членов сообщества – положительное или отрицательное. Он также может объединять мнения, которые имплицитно выражаются в дискуссиях. Предполагается, что профессиональные злоумышленники будут распространять знания и, таким образом, публиковать сообщения с высоким индексом тональности (как положительных, так и отрицательных). Таким образом, главной исследовательской гипотезой является то, что положительные или отрицательные тональности сообщений определяют компетентность злонамеренных пользователей.

Для сбора данных использовалась онлайн-платформа, созданная Университетом Аризоны, на которой размещены собранные коллекции текстовых данных с различных форумов по разным тематическим направлениям. Исходя из целей исследования, были использованы данные с форума "*HackhoundForum*", который содержит 4242 сообщения по различным темам взлома информационных систем и совершения вредоносных действий. Данные являются историческими и датируются с октября 2012 по сентябрь 2015 года.

Для классификации различных ролей злоумышленников в данном случае было решено использовать мультиномиальную логистическую регрессионную модель. В данной работе выделены 5 целевых классов для зависимой переменной и соответствующие им величины независимых переменных, определяющих вхождение пользователя в конкретный класс. Посчитанные средние значения независимых переменных, определяющих вхождение пользователя в конкретную категорию злоумышленника, представлены в таблице 1.

Таблица 1

Средние значения независимых переменных, определяющих вхождение пользователя в конкретную категорию злоумышленника

	β_1 – средняя длина сообщения	β_2 – среднее количество сообщений, опубликованных пользователем	β_3 – среднее общее количество тематических потоков, в которых участвует пользователь	β_4 – среднее количество ключевых слов, относящихся к информационной безопасности	β_5 – средний индекс тональности опубликованного пользователем сообщения
Эксперт=5	1374	145	14	252	0,861
Продвинутый член сообщества=4	928	95	9	197	0,354
Пользователь среднего уровня=3	535	56	5	134	0,134
Пользователь ниже среднего уровня=2	342	32	4	91	0,87
Новичок=1	110	11	1	37	0,25

В рамках независимых переменных ключевыми стали: среднее количество ключевых слов, относящихся к информационной безопасности, а также индекс тональности опубликованного пользователем сообщения. Для расчета данных переменных применялся комплекс взаимосвязанных операций, составление словаря ключевых слов. Словарь составлялся с помощью метода TF-IDF, при котором объединялись частота термина (tf) и обратная частота документа (idf), чтобы получить композитный вес для каждого термина в каждом корпусе для каждого пользователя. Для данных целей использовался нормализованный TF-IDF, вычисляемой по следующей формуле:

$$tf_{i,j} \cdot idf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \cdot \log_2\left(\frac{|D|}{|d|t_i} \in d|\right), \quad (1)$$

где $tf_{i,j}$ является числом вхождений t_i в документ d_i, нормализованный по общему количеству терминов, $idf_{i,j}$ является обратным соотношением документов в рамках t_i и всех документов в корпусе D. Далее применяется мера *Overlap Score Measure* как совокупная сумма баллов *tf-idf* по всем терминам, появляющимся в списке ключевых слов по кибербезопасности, умноженная на количество раз каждый встречаемости ключевых слов кибербезопасности в рамках d. Таким образом, общий вес рассчитывается по формуле:

$$Score = \sum_{t \in k} tf_t - idf_{t,d} . \quad (2)$$

Далее на базе сгенерированного списка ключевых слов, связанных с информационной безопасностью, были созданы списки положительных и отрицательных тональностей, после чего отобранные слова были присоединены к уже существующему списку ключевых слов в рамках программного обеспечения *SentiStrengh*, после чего каждому из слов был присвоен вес (от –5 до 5) в зависимости от близости отношения к теме информационных преступлений. Для автоматизации процесса извлечения ключевых слов из сообщений был настроен алгоритм деревьев решений. В процессе обучения точность данного алгоритма достигла наиболее высокого по сравнению с другими алгоритмами (SVM, k-NN) значения – 78,74%. Для определения интегрального индекса тональности сообщения использовалась следующая формула:

$$TotalSentiment = |PositiveSentiment| + |NegativeSentiment|.$$

Таким образом, в результате работы с данными переменными была создана автоматизированная база для классификации модифицированной лексики злоумышленников.

Далее в таблице 2 приведены полученные коэффициенты (*Coefficient – Coeff.*) и стандартные ошибки (*Standard Error – S.E.*) в рамках модели мультиномиальной логистической регрессии для двух из пяти классов: наиболее высокого (Эксперт) и наиболее низкого (Новичок) по рейтингу.

Таблица 2

Коэффициенты и стандартные ошибки в рамках модели мультиномиальной логистической регрессии для двух из пяти классов

Переменные	Эксперт=5		Новичок=1	
Коэффициенты и стандартные ошибки	*Coeff.*	*S.E.*	*Coeff.*	*S.E.*
средняя длина сообщения: количество букв в каждом сообщении	-0,273	0,134	-0,685	0,112
количество сообщений, опубликованных пользователем	-0,227	0,157	-0,468	0,123
общее количество тематических потоков, в которых участвует пользователь	-0,269	0,311	-0.567	0,054
количество ответов, публикуемых пользователем в каждом тематическом блоке	-0,411	0,146	-1,086	0,056
количество ключевых слов, относящихся к информационной безопасности	-0,132	0,131	-0,538	0,345
индекс тональности опубликованного пользователем сообщения	0,431	0,123	-0,088	0,208
Constant	-5,643	0,346	-2,546	0,173
Observations=99738 *Log likelihood Chi-Sq*=720,61 $p<0.05$				

Таким образом, предлагаемая модель мультиномиального логита построена следующим образом: **Rank of user**$=\beta_0+\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\varepsilon$.

Заключение

Подводя итог работы, следует отметить:

– осуществлен научный обзор существующих теоретических и практических разработок в отечественной и мировой практике в области анализа человеческого фактора в информационной безопасности;

– разработана дорожная карта для анализа форумов как площадок интеграции злонамеренных пользователей;

– создан словарь ключевых слов и их тональностей, определяющий вербальный профиль злоумышленника в области информационной безопасности;

– построена модель мультиномиальной логистической регрессии, в которую включены наиболее важные переменные для классификации злоумышленников;

– выявлена перспективная область изучения человеческого фактора в информационной безопасности на основе интеграции психолингвистики, социологии и статистического моделирования.

Список литературы:

1. Positive Technologies // Актуальные киберугрозы: II квартал 2019 года. URL: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2019-q2/
2. Симанков В.С., Черкасов А.Н. Оптимизация функционирования ситуационного центра при решении задач в различных предметных областях // Естественные и технические науки. 2011. № 4 (54). С. 430–433.
3. Fortune Media IP Limited // Hackett R. Facebook Awards Server-crushing Hacker with Its Biggest Ever Bounty: сайт. URL: http://fortune.com/2017/01/19/facebook-hacker-bug-bounty/
4. Examining Hacker Participation Length in Cybercriminal Internet-Relay-Chat Communities / V. Benjamin, B. Zhang, J.F. Nunamaker, H. Chen // Journal of Management Information Systems. 2016. No. 33. P. 482–510.
5. Benjamin V., Chen H. Developing understanding of hacker language through the use of lexical semantics // In Proceedings of the International Conference on Intelligence and Security Informatics (ISI). 2015. No. 5. P. 79–84.
6. An Analysis of underground forums / M. Motoyama, D. McCoy, K. Levchenko [et al.] // In Proceedings of ACM SIGCOMM Conference on Internet Measures. 2011. No. 2. P. 71.
7. Holt T.J., Kilger M. Know Your Enemy: The Social Dynamics of Hacking // Honeynet Project. 2012. No. 2. P. 1–17.
8. Yip M., Shadbolt N., Webber C. Why Forums? An Empirical Analysis of the Facilitating Factors of Carding Forums // ACM Web Science. 2013. No. 2. P. 26–31.
9. Samtani S., Chinn R., Chen H. Exploring hacker assets in underground forums // In Proceedings of the International Conference on Intelligence and Security Informatics (ISI). 2015. No. 6. P. 31–36.
10. Samtani S., Chen H. Using social network analysis to identify key hackers for keylogging tools in

References:

1. Positive Technologies // Актуальные киберугрозы: II квартал 2019 года. URL: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2019-q2/
2. Simankov V.S., Cherkasov A.N. Optimization of the functioning of the situation center when solving problems in various subject areas // Natural and Technical Sciences. 2011. No. 4 (54). P. 430–433.
3. Fortune Media IP Limited // Hackett R. Facebook Awards Server-crushing Hacker with Its Biggest Ever Bounty: website. URL: http://fortune.com/2017/01/19/facebook-hacker-bug-bounty/
4. Examining Hacker Participation Length in Cybercriminal Internet-Relay-Chat Communities / V Benjamin, B. Zhang, J.F. Nunamaker, H. Chen // Journal of Management Information Systems. 2016. No. 33. P. 482–510.
5. Benjamin V., Chen H. Developing understanding of hacker language through the use of lexical semantics // In Proceedings of the International Conference on Intelligence and Security Informatics (ISI). 2015. No. 5. P. 79–84.
6. An Analysis of underground forums / M. Motoyama, D. McCoy, K. Levchenko [et al.] // In Proceedings of ACM SIGCOMM Conference on Internet Measures. 2011. No. 2. P. 71.
7. Holt T.J., Kilger M. Know Your Enemy: The Social Dynamics of Hacking // Honeynet Project. 2012. No. 2. P. 1–17.
8. Yip M., Shadbolt N., Webber C. Why Forums? An Empirical Analysis of the Facilitating Factors of Carding Forums // ACM Web Science. 2013. No. 2. P. 26–31.
9. Samtani S., Chinn R., Chen H. Exploring hacker assets in underground forums // In Proceedings of the International Conference on Intelligence and Security Informatics (ISI). 2015. No. 6. P. 31–36.
10. Samtani S., Chen H. Using social network analysis to identify key hackers for keylogging tools in

hacker forums // In Proceedings of the International Conference on Intelligence & Security Informatics (ISI). 2016. No. 2. P. 319–321.

11. Moving toward black hat research in information systems security: An editorial introduction to the special issue / A.M. Mahmood, M. Siponen, D. Straub [et al.] // MIS Quarterly. 2010. No. 34. P. 2–22.
12. Benjamin V., Chen H. Securing cyberspace: Identifying key actors in hacker communities // Proceedings of the International Conference on Intelligence and Security Informatics (ISI). 2012. No. 3. P. 24–29.
13. Social network analysis of a criminal hacker community / Y. Lu, X. Luo, M. Polgar, Y. Cao // Journal of Computer Information Systems. 2010. No. 51. P. 31–41.
14. Examining the social networks of malware writers and hackers / T.J. Holt, D. Strumsky, O. Smirnova, M. Kilger // International Journal of Cyber Criminology. 2012. No. 6. P. 891.
15. Hacker J., Bernsmann R., Riemer K. Dimensions of User Behavior in Enterprise Social Networks. Social Knowledge Management in Action. Springer International Publishing, 2017. No. 11. P. 125–146.

hacker forums // In Proceedings of the International Conference on Intelligence & Security Informatics (ISI). 2016. No. 2. P. 319–321.

11. Moving toward black hat research in information systems security: An editorial introduction to the special issue / A.M. Mahmood, M. Siponen, D. Straub [et al.] // MIS Quarterly. 2010. No. 34. P. 2–22.
12. Benjamin V., Chen H. Securing cyberspace: Identifying key actors in hacker communities // Proceedings of the International Conference on Intelligence and Security Informatics (ISI). 2012. No. 3. P. 24–29.
13. Social network analysis of a criminal hacker community / Y. Lu, X. Luo, M. Polgar, Y. Cao // Journal of Computer Information Systems. 2010. No. 51. P. 31–41.
14. Examining the social networks of malware writers and hackers / T.J. Holt, D. Strumsky, O. Smirnova, M. Kilger // International Journal of Cyber Criminology. 2012. No. 6. P. 891.
15. Hacker J., Bernsmann R., Riemer K. Dimensions of User Behavior in Enterprise Social Networks. Social Knowledge Management in Action. Springer International Publishing, 2017. No. 11. P. 125–146.