УДК 004.056.53 + 004.4.056 Вестник УрФО № 1(27) / 2018, с. 65–70

Муравьев Н. С., Астахова Л. В.

ПРОФИЛАКТИКА ИНЦИДЕНТОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ПРОФИЛИРОВАНИЯ ПОЛЬЗОВАТЕЛЕЙ: ПРОГРАММНО-ТЕХНИЧЕСКИЙ АСПЕКТ

В статье рассмотрена проблема профилирования пользователей информационной системы как средства профилактики инцидентов в сфере информационной безопасности (ИБ). Предложены направления для технической реализации профилирования. Обоснованы критерии оценки поведения пользователей информационной системы (ИС), имеющие количественную характеристику. Определена возможность автоматизации профилирования и прогнозирования возможных инцидентов ИБ по вине каждого отдельно взятого пользователя ИС. Предложен алгоритм прогнозирования тенденции в поведении пользователя. В качестве инструмента определения степени уязвимости пользователя обоснован «Профайл безопасного пользователя».

Ключевые слова*: информационная безопасность, поведение пользователя, профайлинг, профилирование, профайл, инциденты, пользователь, машинное обучение, прогнозирование поведения.*

Muravyov N. S., Astakhova L. V.

PREVENTION OF INFORMATION SECURITY INCIDENTS BASED ON USER PROFILING: PROGRAM-TECHNICAL ASPECT

The article considers the problem of profiling users of the information system as a means of preventing incidents in the field of information security. Suggested directions for technical implementation of profiling. The criteria for assessing the behavior of IP users, which have a quantitative characteristic, are substantiated. The possibility of automation of profiling and forecasting of possible IS incidents through the fault of each individual IP user has been determined. An algorithm for predicting the trend in user behavior is proposed. The «Safe User Profile» is justified as a tool for determining the vulnerability of a user.

Keywords: *information security, user behavior, profiling, profiling, profile, incidents, user, machine learning, behavior prediction.*

The work was supported by Act 211 Government of the Russian Federation, contract № 02. A03.21.0011

Человеку отводится ключевая роль в процессах организации, функционирования и развития деятельности коммерческих и государственных организаций, предприятий и иных структур. В процессе деятельности сотрудник располагает совокупностью информационных ресурсов организации, а также имеет доступ к средствам ее обработки, которые могут вызвать интерес не только у правообладателя, но и третьих лиц. В целях сохранности ключевой информации в организациях вводится режим информационной безопасности (далее – ИБ), который состоит из правовых, организационных и программно-технических мер защиты.

Большая часть выявленных нарушений информационной безопасности происходит из-за человеческого фактора в информационной системе (далее – ИС). Это подтверждают статистические исследования компании InfoWatch. Например, за 2016 год по причине внутреннего нарушителя было реализовано около 62 % угроз [5]. Под человеческим фактором мы понимаем набор преднамеренных или неумышленных (например, при неправильной эксплуатации ИС) действий внутреннего пользователя, в результате которых реализуется угроза и происходит утечка информации.

В повседневной рабочей среде при длительном периоде без инцидентов в ИБ, а также при отсутствии видимых угроз у пользователя меняется восприятие опасностей. Это приводит к успешной реализации социальной инженерии, фишинга, инсайдерских утечек и т. д., которые предполагают наличие человеческого воздействия на систему. Подобные угрозы усложняются тем, что внутренний пользователь имеет прямой доступ к информации и его поведение с меньшей вероятностью отличается от нормы. Поэтому в последние годы повысился интерес к оценке кадровых уязвимостей информационных систем [1], в том числе на основе профайлинга [6].

Анализ российской и зарубежной литературы показал, что зачастую профайлинг рассматривается с позиций гуманитарно-психологического подхода: в деятельности кадровых служб [2], в правоохранительной деятельности [3] и др. Профайлингу в сфере ИБ уделяется значительно меньше внимания. Этим обусловлена цель данной работы – обосновать программно-техническое решение профилирования пользователей как средство профилактики инцидентов информационной безопасности. В числе задач статьи – определение критериев для оценки поведения пользователя при взаимодействии с ИС для последующего прогнозирования тенденции его поведения; обоснование алгоритма прогнозирования тенденции ИБ-поведения пользователя на основе предложенного «Профайла безопасного пользователя».

Зарубежные специалисты рассматривают профайлинг и профилирование с позиции, основанной на криминалистических подходах [4]. На наш взгляд, профилирование – процесс сбора и накопления данных о пользователях, структурированных по атрибутам согласно определенным критериям оценки с целью прогнозирования поведения пользователей. Профайл – это результат профилирования: совокупность записей в базе данных с атрибутами, содержащими критерии для оценки поведения по каждому пользователю ИС.

Достижение обозначенной цели возможно путем смешения социальных и технологических решений, результаты которых должны вноситься в профайл каждого отдельного взятого сотрудника организации.

К социальным решениям относятся такие данные, как сведения о психологических, социальных, культурных и иных гранях жизни человека. Данные об этом собираются через опросы, тесты, наблюдения и иные формы внутренних и внешних проверок. К технологическому решению относится раздел, который исследует непосредственно поведение человека в момент, когда он является пользователем ИС [4].

Мы рассмотрим идею программно-технического профилирования пользователя как дополнение к различным системам мониторинга или DLP-системам. В совокупности эти данные вносятся в профайл для определения поведенческого типа каждого сотрудника. При анализе данных из профайла изучается поведение сотрудника, соответствие степени соблюдения политики информационной без-

опасности организации, уровень лояльности и прочие поведенческие качества для дальнейшего прогнозирования.

Для разработки программно-технических средств профилирования необходимо четкое определение критериев. Зарубежные специалисты выделяют такие критерии оценки поведения пользователя, как:

1) обращение с паролями (оценка изменения пароля и его надежность в соответствии с установленной политикой безопасности);

2) периодичность резервного копирования (своевременность резервного копирования в случае наличия таких обязанностей на пользователе [4].

Мы считаем, что перечень этих критериев может быть расширен на основе используемых сведений о пользователях, собираемых в DLP-системах и других системах с возможностью мониторинга, применяемых в российской и зарубежной практике защиты информации. К таким критериям мы относим:

- данные о попытках доступа к сектору информации ограниченного доступа для пользователя;
- данные о сетевой активности и сопоставление просматриваемых интернет-ресурсов на предмет отношения к профессиональной деятельности;
- сведения о попытках взаимодействия со съемными носителями;
- регистрация запросов к установленным средствам защиты информации (например, данные о попытке просмотра настроек средства от НСД, антивируса и т. д.);
- количество инцидентов, аномалий и т. д., зафиксированных DLP-системой;
- иные сведения, которые могут быть получены в процессе мониторинга.

Указанный перечень критериев может быть расширен в соответствии с технологическими возможностями каждой отдельной организации.

На основе указанных критериев предлагаем сформировать так называемый «Профайл безопасного пользователя». В нем должны содержаться оптимальные показатели указанных критериев, которые соответствуют установленным требованиям Политики ИБ. Например, такими показателями могут быть установленная периодичность смены паролей пользователей, количество попыток несанкционированного доступа к защищаемой информации и др.

Представим общую схему профилирования, оценки и прогнозирования поведения пользователя ИС (рис. 1).

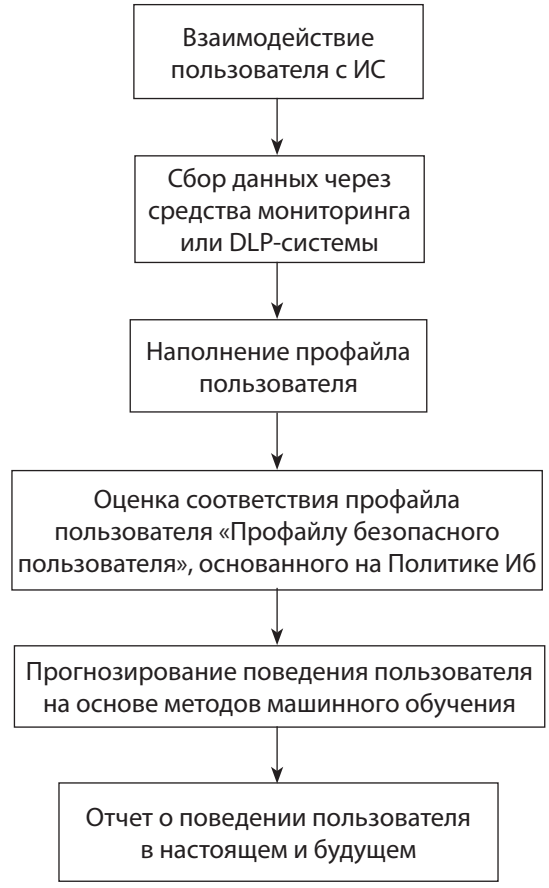

Рис. 1 Общая схема профилирования, оценки и прогнозирования поведения пользователя ИС

Более подробно рассмотрим процесс прогнозирования. Для этого предлагаем применить методы машинного обучения и считать полученную программу частью «Профиля безопасного пользователя», а полученные показатели – учитывать при дальнейшей оценке пользователя.

Машинное обучение – моделирование признаков данных и соответствующих данным меток. Данное определение относится к категории задач «обучение с учителем», когда некоторые прогнозируемые данные частично известны [7].

Сформулируем задачу: на основе данных за два месяца работы пользователя спрогнозировать уровень его угрозы ИБ на предстоящий месяц с помощью компонента для прогнозирования «Профиля безопасного пользователя».

Для решения задачи применяются библиотеки:

• Pandas,
• Numpy,
• Scikit-Learn на языкеPython.

Уровень угрозы рассчитывается как:

$$\text{Уровень угрозы(за месяц)} =$$

$$= \frac{\text{Количество угроз за месяц}}{\text{Кличество дней в месяце}} * 100$$

Собранные данные поступают в программу для прогнозирования в следующем виде (рис. 2):

Дата	Инциденты	Аномалии	Опоздания	Отсутствие	Статус
01.09.2017	0	0	0	0	0
02.09.2017	1	0	0	0	1
03.09.2017	0	0	0	1	0
04.09.2017	0	0	0	0	0
05.09.2017	0	0	0	0	0
06.09.2017	1	1	0	1	1
07.09.2017	0	0	0	0	0
08.09.2017	0	0	0	0	0
09.09.2017	2	0	1	0	1

Рис. 2. Пример части данных для прогнозирования

Представленные данные необходимо разделить на матрицу признаков (X) и целевой вектор (Y).

Матрица признаков – набор столбцов: инциденты, аномалии, опоздания, отсутствие на рабочем месте (заранее выбранные критерии для прогнозирования).

Целевым вектором является столбец «Статус», значения которого и будут спрогнозированы, где 0 – пользователь не являлся угрозой, 1 – пользователь являлся угрозой. В данном случае пользователя считали угрозой ИБ, если у него был, хотя бы один инцидент.

На основе уже заполненного целевого вектора за сентябрь и октябрь рассчитаем для них уровень угрозы пользователя:

• Сентябрь

```
In [165]: print (september_result,"% уровень угрозы за сентябрь")
18 % уровень угрозы за сентябрь
```

• Октябрь

```
In [168]: print (october_result,"% уровень угрозы за октябрь")
13 % уровень угрозы за октябрь
```

Алгоритм прогнозирования представим в виде следующих этапов (далее указаны только основные функции программы):

1. Выбор модели: «Гауссов наивный байесовский классификатор»:
from sklearn.naive_bayes import GaussianNB;

Выбранная модель «Гауссов наивный байесовский классификатор» – модель из категории «обучение с учителем», т. е. целевой массив изначально уже имеет значения, на основе которых прогнозируются значения на будущий месяц. Данные классифицируются относительно целевого вектора «Статус».

2. Создание экземпляра модели:
model = GaussianNB()
3. Обучение модели:
model.fit(X_dlp, Y_dlp)
4. Предсказание значений:
november = model.predict(X_dlp)
5. Расчет и вывод результата:
november_result =
round((sum/november.size)*100)

```
In [62]: print ("Прогнозируем", november_result,"% уровень угрозы за ноябрь")
Прогнозируем 23 % уровень угрозы за ноябрь
```

Мы видим, что прогноз уровня угрозы ИБ со стороны данного пользователя на месяц ноябрь составляет 23 %. А это значит, что прогнозируется повышение тенденции на 10 % относительно октября. Это можно расценивать как повышение угрозы со стороны данного пользователя.

Таким образом, важным средством профилактики инцидентов ИБ является профилирование пользователей. Обоснованные критерии оценки поведения пользователей ИС имеют количественную характеристику. Это обеспечивает возможность автоматизации профилирования и прогнозирования возможных инцидентов ИБ по вине каждого отдельно взятого пользователя ИС. В статье обоснован алгоритм прогнозирования тенденции в ИБ – поведении пользователя, под которой понимается разница в показателе оценки уровня угрозы пользователя на выбранный прогнозируемый период. В качестве инструмента определения степени уязвимости пользователя предложен «Профайл безопасного пользователя».

Литература

1. Астахова Л. В. и др. Автоматизация многофакторной оценки кадровых уязвимостей информационной безопасности / Л. В. Астахова, В. А. Ефремов, А. И. Митькин // Вестник УрФО. Безопасность в информационной сфере. – 2014. - № 4 (14). – С. 57–61.

2. Арпентьева М. Р. Профайлинг как современная HR-технология // Инновационное развитие современных социально-экономических систем материалы III Международной заочной научно-практической конференции. – 2016. – С. 296–301.

3. Черкасова Е.С. Профайлинг как метод создания психологического портрета потенциального преступника на этапе организации предварительного расследования // Юридическая наука и практика. – 2013. – С. 72–75.

4. Fernando S. A. Securing Information Sharing Through User Security Behavioral Profiling / S. A. Fernando, Takashi Yukawa // Transactions on Engineering Technologies. – 2013. – С. 655–670.

5. Глобальное исследование утечек конфиденциальной информации в 2016 году. – URL: https://www.infowatch.ru/analytics/reports/17479 (дата обращения: 10.01.2018).

6. Бируля И. Как оцифровать человеческий фактор // Журнал «Information Security/ Информационная безопасность». – 2017. – № 4. – URL: http://www.itsec.ru/articles2/sys_ogr_dost/kak-otsifrovat-chelovecheskiy-faktor

7. Плас Дж. Вандер. Python для сложных задач: наука о данных и машинное обучение. — СПб. : Питер, 2018. – 576 с.

Referenses

1. Astahova L.V. i dr. Avtomatizaciya mnogofaktornoj ocenki kadrovyh uyazvimostej informacionnoj bezopasnosti / L. V. Astahova, V.A. Efremov, A.I. Mit'kin // Vestnik UrFO. Bezopasnost' v informacionnoj sfere. – 2014. – № 4 (14). – S. 5–61.

2. Arpent'eva M. R. Profajling kak sovremennaya HR-tekhnologiya // Innovacionnoe razvitie sovremennyh social'no-ehkonomicheskih sistem materialy III Mezhdunarodnoj zaochnoj nauchno-prakticheskoj konferencii. -2016.- S. 296-301.

3. CHerkasova E.S. Profajling kak metod sozdaniya psihologicheskogo portreta potencial'nogo prestupnika na ehtape organizacii predvaritel'nogo rassledovaniya// YUridicheskaya nauka i praktika.-2013.-S.72-75.

4. Fernando S. A. Securing Information Sharing Through User Security Behavioral Profiling / S. A. Fernando, Takashi Yukawa // Transactions on Engineering Technologies. – 2013. – P. 655–670.

5. Global'noe issledovanie utechek konfidencial'noj informacii v 2016 godu – https://www.infowatch.ru/analytics/reports/17479 – (data obrashcheniya: 10.01.2018).

6. Birulya I. Kak ocifrovat' chelovecheskij faktor // ZHurnal "Information Security/ Informacionnaya bezopasnost'". – 2017. – № 4. http://www.itsec.ru/articles2/sys_ogr_dost/kak-otsifrovat-chelovecheskiy-faktor

7. Plas Dzh. Vander. Python dlya slozhnyh zadach: nauka o dannyh i mashinnoe obuchenie. — SPb.: Piter, 2018. — 576 s.

МУРАВЬЕВ Никита Сергеевич, студент кафедры защиты информации Южно-Уральского государственного университета. 454080, г. Челябинск, пр. Ленина, д. 76. E-mail: mns7496@yandex.ru

АСТАХОВА Людмила Викторовна, доктор педагогических наук, профессор, профессор кафедры защиты информации Южно-Уральского государственного университета. 454080, г. Челябинск, пр. Ленина, д. 76. E-mail: lvastachova@mail.ru

MURAVYOV Nikita, student of the Department of Information Security South Ural State University. 454080, Chelyabinsk, Lenin Avenue, 76. E-mail: mns7496@yandex.ru

ASTAKHOVA Lyudmila, doctor of pedagogical sciences, professor, professor of the department of information protection South Ural State University. 454080, Chelyabinsk, Lenin Avenue, 76. E-mail: lvastachova@mail.ru