Современное состояние утечек конфиденциальной информации через облачные хранилища данных

Д.Д. Дайнеко[1], Д.А. Бахтеева[1], Д.Г. Зыбин[1], В.А. Спирин[1], А.В. Калач[1,2]

[1]Воронежский институт Федеральной службы исполнения наказаний, Иркутская 1-а, Воронеж, Россия, 394072
[2]Воронежский государственный технический университет, Московский проспект 14, Воронеж, Россия, 394026

Аннотация. Представлена информация об утечках конфиденциальных данных, находящихся в облачных хранилищах типа Amazon, Mongo DB, в файловых хостингах типа Google Drive, а также облачных серверов резервного копирования в период 2016-2018 гг. Сделаны выводы о необходимости развития систем обеспечения безопасности информации с использованием облачных технологий за счет повышения их киберустойчивости.

1. Введение

Все большую значимость на рынке информационных ресурсов приобретают облачные технологии. Государственные организации и коммерческие компании все больше ощущают удобство использования облачных сервисов в виртуальных средах и все чаще готовы выносить туда свои базы данных. По мере развития как самих облачных технологий, так и информационной среды в целом, случаи непреднамеренной компрометации данных встречались эпизодически. Но в настоящее время угроза потери данных становится все более ощутимой и количество инцидентов регистрируется все чаще. Большая часть всех утечек конфиденциальной информации из открытых хранилищ есть следствие неправильного предоставления прав доступа лицам или же неверная настройка конфигурации систем [1 – 8].

Специалистами аналитического центра InfoWatch было проведено изучение случаев утечек корпоративных баз данных через облачные сервисы и другие файловые хранилища. Предметом стали различные публичные сообщения об утечках конфиденциальных данных через облачные хранилища (Amazon, Mongo DB и др.), папок в файловых хостингах (Google Drive и др.), а также облачных серверов резервного копирования в период 2016-2018 гг. При этом, не рассматривались инциденты, связанные с утечками информации через веб-серверы и почтовые серверы. Акцент был сделан на проблемы утечек персональных и иных данных из объектах виртуализации данных (облачные хранилища и сервисы) [9].

2. Утечка конфиденциальной информации через облачные хранилища данных

На основе публичных сообщений в СМИ и других открытых источников аналитический центр InfoWatch в 2018 году зафиксировал увеличение утечек конфиденциальных данных через облачные серверы и хранилища с доступом через Сеть почти в 1,5 раза по сравнению с предыдущим годом. И такая тенденция сохраняется на протяжении нескольких лет. Например, по сравнению с 2016 г. утечек стало больше в 4,4 раза [9].

Обладателем печального «рекорда» по числу скомпрометированных записей является 2017 год. Более 1,7 млрд записей было потеряно с незащищенных серверов (а это около 13% всего объема потерянных в этом году данных). Большая часть данных принадлежала компании, занимающейся сетевым маркетингом – River City Media. Вследствие неправильного резервного копирования была допущена ошибка и огромная база данных объемом 1,34 млрд записей была скомпрометирована. За 2018 год потери данных из открытых серверов составляют 1,3 млрд записей, в самом крупном инциденте было потеряно 400 млн записей. При этом, порядка 40% утечек приходится на высокотехнологичные компании

На период 2017 - 2018 гг., наибольшая доля утечек, связанных с неправильным обслуживанием и настройкой облачных серверов, а также других ошибок при работе с виртуализацией данных, приходится на высокотехнологичный сегмент – производители ИТ-продукции, ИТ-сервисы, социальные сети и т.д. [9].

Высокоразвитые компании, обладающие технологическим потенциалом склонны к использованию трендовых инструментов визуализации, а в частности – внешних хранилищ данных. Но к сожалению, не всегда сотрудники таких компаний учитывают некоторые нюансы при работе с облачными серверами, и как следствие, учащаются случаи утечек.

Таким образом, в результате около 90% всех данных, которые были утеряны с незащищенных серверов в 2018 г., пришлись именно на хайтек-индустрию (таблица 1).

Таблица 1. Отраслевое распределение утечек из открытых облачных ресурсов за период 2017-2018 гг.

Отрасли утечек	2017 г.	2018 г.
Государственные организации	14,3%	8,5%
Медицина	8,2%	11,4%
Высокие технологии	32,6%	40%
Образование	4,1%	7,2%
Ритейл	10,2%	7,2%
Промышленность и транспорт	12,2%	7,2%
Банки, финансы и страхование	12,2%	1,4%
Другое	6,2%	17,1%

На протяжении того же года в число «жертв» стали включаться сервера медицинских и образовательных учреждений – количество утечек сильно возросло. Но в то же время значительно падает количество потери данных из финансовой сферы, промышленности и госсектора. Персональные данные утекают в четырех случаях из пяти. Большая часть, а именно более 80% утечек приходится на персональные данные. По 9,2% случаев компрометации данных с открытых хранилищ – это утечки платежной информации, а также коммерческих секретов и производственных ноу-хау (таблица 2) [9].

Таблица 2. Распределение утечек по типам данных за период 2017-2018 гг.

Типы данных	2017 г.	2018 г.
Персональные данные	77,8%	81,6%
Коммерческие секреты	8,9%	9,2%
Платежная информация	6,7%	9,2%
Государственная тайна	6,7%	-

Четверть утечек происходит с серверов Amazon S3. В части, касающейся распределения скомпрометированных серверов по типам произошли существенные изменения. Лидером в 2018 году, как и в предыдущие было облачное хранилище Amazon S3. На их долю пришлось более четверти утечек (таблица 3).

Тенденцией к утечкам конфиденциальной информации в 2018 году обзавелись сервера Mongo DB, а также такие платформы как Elasticsearch и Apache, файловый хостинг Google

Drive. В свою очередь такие сервисы как rsync и GitHub сократили случаи утечки данных при резервном копировании в 3 и 7 раз соответственно. В каждом втором инциденте жертвой становится американская компания. Что же касается распределения утечек по странам – тут произошли сильные изменения. 2018 г. как и в предыдущие лидером остается США , но утечки тут сократились с 75,5% до 47,1%. А вот в Канаде количество увеличилось почти в 3 раза, в Индии – в 2 раза (таблица 4) [9].

Таблица 3. Данные о скомпрометированных хранилищах, 2017-2018 гг.

Хранилища	2017 г.	2018 г.
Amazon S3	28.6%	25.7%
Apache	2%	4.3%
FTP	2%	1.4%
GitHub	10.2%	1.4%
Microsoft SQL	8.2%	1.4%
Mongo DB	6.1%	15.7%
rsync	8.2%	2.8%
Другое	34.7%	34.3%
Google Drive	-	5.7%
Elasticsearch	-	7.1%

Таблица 4. Распределение утечек по странам, 2017-2018 гг.

Страны	2017 г.	2018 г.
Канада	2,0%	5,9%
Китай	2,0%	1,5%
Великобритания	6,1%	4,4%
Индия	4,1%	8,8%
Швеция	2,0%	-
США	75,5%	47,1%
Другое	8,2%	19,1%
Финляндия	-	2,9%
Франция	-	2,9%
Россия	-	2,9%
Бразилия	-	4,4%

Утечки данных из незащищенных облачных хранилищ приобрели угрожающий масштаб во всем мире. Многие крупные компании, такие как American Express, Honda, Nokia, Sky Brazil, и ряд государственных структур стали жертвами таких инцидентов. Потенциальным источником утечки данных является каждое хранилище, в котором периодически выявляют тысячи и даже десятки тысяч ошибок в конфигурации.

Одна ошибка в работе корпорации может являться причиной утечки с незащищенного сервера. Многие компании теряют огромные базы данных, так как представители киберпреступности проводят мониторинг облачных ресурсов.

Проведенный анализ данных об утечках конфиденциальной информации через облачные хранилища данных позволяет сделать вывод о целесообразности повышения их киберустойчивости. Киберустойчивость представляет собой частный случай одностороннего информационного конфликта в киберпространстве, в котором атакующая сторона для достижения своих целей использует различные стратегии, а обороняющаяся – стратегии обеспечения устойчивого функционирования системы управления объектами защиты от подобных воздействий. При этом, особенностью кибернетического противоборства заключается в том, что как минимум две (или более) подсистемы управления стремятся распространить управляющие воздействие друг на друга через совместно используемый общий ресурс (глобального информационного пространство) [10]. Вследствие этого в качестве одного

из исходных положений можно рассматривать гипотезу о поведении информационных систем в условиях в условиях информационного конфликта, фиксирующая множество возможных стратегий нападения и защиты. На этапе формализации процесса киберустойчивости система гипотез порождает соответствующее разнообразие моделей и алгоритмов киберустойчивости.

Концептуальные модели киберустойчивости объектов защиты в условиях информационного конфликта фиксируют множество возможных априорных и апостериорных знаний о стратегиях противоборствующих сторон и выступают в качестве их вербальных (слабоформализованных) моделей (рисунок 1).

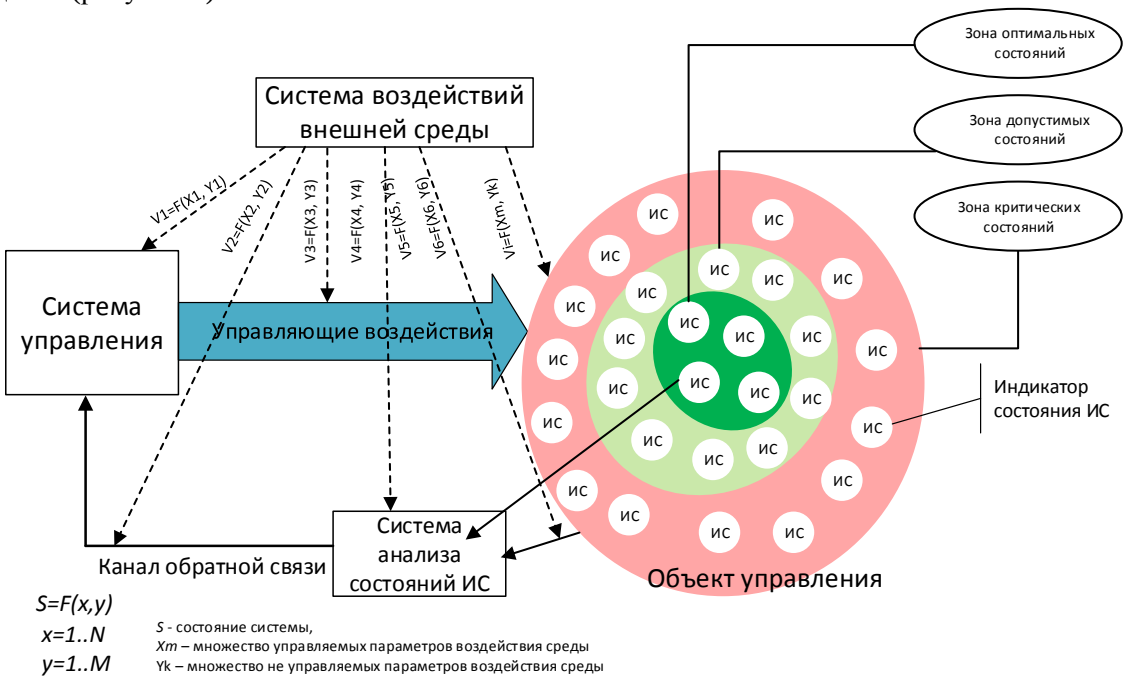

Рисунок 1. Графическая модель представления многовариантного состояния объекта исследования в условия воздействия.

Формально модель проблемной ситуации может быть задана с помощью следующих элементов, отражающих множества наиболее существенных с точки зрения ЛПР факторов [10]:

$N = \{1,2,...,i_0\}$ – множество участков конфликта;

$S \subseteq N$ – подмножество участков конфликта, объединенных общей целью, содействующих друг с другом (коалиция);

R – отношение соподчиненности на коалициях;

U^S – множество стратегий S-ой коалиции;

Θ^S – информация, имеющаяся в распоряжении S-ой коалиции;

$\underset{S}{X} U^S$ – декартово произведение множеств стратегий коалиций – множество ситуаций;

$W^S\left(\underset{S}{X} U^S\right)$ – функция выигрыша S-ой коалиции;

P^S – модель, отражающая предпочтения S-ой коалиции на множестве ситуаций в модели проблемной ситуации.

Для каждой коалиции на множестве ситуаций в общем случае необходимо задать бинарное отношение предпочтения. Целью коалиции в конфликте является достижение наиболее предпочтительной в том или ином смысле ситуации. Ситуацией называется результат выбора всеми игроками своих стратегий. В общем случае выигрыш i-го игрока не совпадает с

выигрышем $W^{\{i\}}(\cdot)$, который он может себе обеспечить, действуя в одиночку, так как, вступив в коалицию S, он может получить больше. Это объясняется тем, что:

$$W^S(\cdot) \geq \sum_{i \in S} W^{\{i\}}(\cdot).$$

Таким образом, наиболее общую модель конфликта задает система множеств:

$$\left\langle N, \left\{U^S, \theta^S, \mathrm{P}^S, W^S\right\}_{S \subset N}, R\right\rangle.$$

Механизм H модели конфликта состоит в том, что участки конфликта из множества N оказывают воздействие на некоторую систему, что в итоге приводит к получению ими определенных выигрышей:

$$H : U^{\{1\}} \times U^{\{2\}} \times \cdots \times U^{\{i_0\}} \times T \to \left(W^{\{1\}}, ..., W^{\{i_0\}}\right),$$

где T – множество моментов времени развития конфликта. Каждый из участков конфликта действует по вполне определенным правилам, стремясь достичь своей цели. Предполагается, что все участки конфликта получают какую-либо информацию $\theta^{\{i\}}$ о состоянии системы. В большинстве случаев информация $\theta^{\{i\}}$ касается вида функций $W^{\{i\}}$ выигрыша отдельных игроков или W^S их коалиций, а также множества допустимых стратегий участков конфликта и коалиций.

3. Заключение

Таким образом, помимо повышения уровня квалификации системных администраторов и пользователей, следует проводить ревизии информационных ресурсов, использовать инструменты контроля доступа, проводить мониторинг незащищенных облачных хранилищ информации ограниченного доступа и активно применять современные модели киберустойчивости.

4. Литература

[1] Глотина, И.М. Латентный характер угроз экономической безопасности / И.М. Глотина // Материалы III международной научно-практической конференции. – Саратов: ООО “Центр профессионального менеджмента “Академия Бизнеса”, 2015. – С. 51-56.

[2] Шпеко, М.В. Шифрование данных в облачных сервисах / М.В. Шпеко // Материалы Всероссийской молодежной научно-практической школы. – Кемерово: Кузбасский государственный технический университет им. Т.Ф. Горбачева, 2014. – С. 281-282.

[3] Батурина, М.В. Будущее за киберстрахованием / М.В. Батурина // Страховое дело, 2018. – Т. 7, № 304. – С. 28-33.

[4] Пригодин, С.А. Правовая основа кибербезопасности в Российской Федерации и тенденции развития / С.А. Пригодин // Сборник научных трудов по материалам VII Международной научно-практической конференции – Анапа: ООО “Научно-исследовательский центр экономических и социальных процессов”, 2019. – С. 36-41.

[5] Бобрышева, Г.В. Облачная безопасность / Г.В. Бобрышева, Е.М. Пудовкина // Сборник статей XVII Международной научно-технической конференции – Пенза: Автономная некоммерческая научно-методическая организация “Приволжский Дом знаний”, 2017. – С. 99-104.

[6] Калач, А.В. Системный анализ и оценка современных угроз обеспечения безопасности информации / А.В. Калач // Вестник Воронежского института ФСИН России. – 2019. – № 1. – С. 69-74.

[7] Гребенников, Н. Киберугрозы сегодня: предупрежден – значит, вооружен / Н. Гребенников // Первая миля. – 2017. – Т. 4, № 65. – С. 76-78.

[8] Полтавцева, М.А. Консистентный подход к построению защищенных систем обработки и хранения больших данных / М.А. Полтавцева // Проблемы информационной безопасности. Компьютерные системы. – 2019. – № 2. – С. 29-44.

[9] Аналитический центр InfoWatch [Электронный ресурс]. – Режим доступа: https://www.infowatch.ru/ (31.10.2019).
[10] Комарович, В. Ф. Компьютерные информационные войны: концепция и реалии / В.Ф. Комарович, И.Б. Саенко // Защита информации. Конфидент. – 2002. – № 4-5. – С. 84-88.

The current status of sensitive information leaks through cloud storage

D.D. Dayneko[1], D.A. Bakhteeva[1], D.G. Zybin[1], VA. Spirin[1], A.V. Kalach[1,2]

[1]VRI of the FPS of Russia, Irkutskaya str. 1-a, Voronezh, Russia, 394072
[2]VSTU, 20-letya Oktyabry str. 84, Voronezh, Russia, 394026

Abstract. Provides information about leaks of sensitive data stored in cloud storage like Amazon, Mongo DB, file hosts like Google Drive, as well as cloud backup servers in the period 2016-2018. Conclusions drawn on the need to develop information security systems using cloud technologies by increasing their cyber stability.