

## О связи утечки персональных данных с показателями кибербезопасности

*И.Н. Карпенко*

*Финансовый университет при Правительстве РФ*

**Аннотация:** Несмотря на меры правительств стран по защите данных, тем не менее наблюдается рост их утечек. В исследовании выдвинуты 2 гипотезы. 1-я - об обратной связи плотности взломанных аккаунтов на 1000 человек и индексов кибербезопасности GCI и NCSI. На 1 этапе эта гипотеза была опровергнута. 2-я гипотеза о существовании факторов, взаимодействующих с факторами кибербезопасности и влияющих на утечку персональных данных, была подтверждена. Выявленным фактором явился показатель ВВП на душу населения. После элиминации влияния этого фактора, корреляционный анализ подтвердил наличие обратной связи между показателем утечки персональных данных в форме показателя плотности и индексами кибербезопасности GCI и NCSI. Корреляционные коэффициенты составили -0,43 и -0,48, соответственно.

**Ключевые слова:** кибернарушение, кибербезопасность, GCI, NCSI, утечка данных, взлом аккаунтов, персональные данные, конфиденциальная информация, анализ данных, корреляция.

Актуальность исследования обоснована ростом нарушений в сфере безопасности информации прежде всего персональных данных (далее ПД) в глобальной сети. Это приводит к материальному и моральному ущербу для населения разных стран мира. Предпринимаемые рядовыми пользователями меры не могут обеспечить полную защиту их персональных данных. Для защиты конфиденциальных данных, в том числе ПД своих граждан, правительства разных стран проводят политику, направленную на обеспечение кибербезопасности (далее КБ). Использование киберпреступниками самых современных цифровых технологий требует постоянного совершенствования всего набора мер противодействия: правовых, организационных, технических, межстранового сотрудничества.

Однако анализ публикуемых данных показывает, что в последние годы, несмотря на предпринимаемые правительствами стран усилия по укреплению КБ, наблюдается рост количества утечек ПД. Это, например,

---

отмечено в аналитическом отчете компании сферы информационной безопасности InfoWatch (Россия, Москва) «Утечки информации в мире, 2022-2023 годы». В нем отражен рост количества утечек данных в целом в мире в 2023 году по сравнению с 2022 годом более, чем на 60%, а ПД – более, чем в 2 раза [1].

На основе анализа наблюдений встает вопрос о значимости связи утечки ПД и реализуемых мер КБ. Это станет предметом нашего исследования.

В начале следует отметить, что в настоящее время нет единого подхода к определению понятий «кибербезопасность», «утечка данных», «персональные данные». Так, большинство российских исследователей считают, что информационная безопасность более широкое понятие, чем КБ. Например, Ю.А. Баландин на основе анализа нормативных правовых актов России делает вывод, что КБ – один из элементов информационной безопасности, характеризующийся применением программно-аппаратных средств [2].

Зарубежные исследователи чаще имеют противоположную точку зрения. К ним относятся и взгляды зарубежных компаний, занимающихся разработкой программного обеспечения для защиты конфиденциальной информации пользователей. Для нашей работы возьмем определение понятий, представленных на сайте корпорации International Business Machines Corporation (The United States, Armonk) [3]. КБ – это любая технология, мера или практика для предотвращения кибератак и смягчения последствий вредоносных взломов. Она включает безопасность искусственного интеллекта, критической инфраструктуры, конечных точек, приложений; сетевую, облачную, информационную, мобильную типы безопасности. ПД – любая информация, которая может использоваться для

---

раскрытия личности конкретного лица, в том числе адрес электронной почты.

Выдвинем гипотезу о связи страновых показателей утечки ПД и КБ.

Гипотеза  $H_0$ . Существует значимая обратная связь между значениями  $Y$  (страновые утечки ПД) и уровнем фактора  $A$  (КБ), т.е. страновое уменьшение величины утечек ПД значимо связано с увеличением уровня КБ.

Гипотеза  $H_{01}$ . Существует взаимодействие фактора  $A$  (КБ) и фактора  $B$  (не установлен), влияющее на значения  $Y$  (страновые утечки ПД).

В качестве  $Y$ , характеризующего страновые утечки ПД, возьмем плотность взломанных адресов электронной почты (аккаунтов), используемых для регистрации в онлайн-сервисах, на 1000 человек населения (далее Пл).

В качестве показателей КБ используем индексы Global Cybersecurity Index (GCI) и National Cybersecurity Index (NCSI).

Для проведения исследования возьмем данные о взломанных аккаунтах за 2023 год из таблицы «Data breach statistics by country: 2023vs2022» [4] компании по КБ Surfshark (The Netherlands, Amsterdam).

Для дальнейшего анализа произведем следующую подготовку данных:

1. Из выборки 250 стран исключим страны с численностью населения менее 1 млн человек. Низкая численность проживающих в таких странах приводит к недостоверности результатов их статистической обработки.

2. Произведем расчет Пл. В качестве источника данных о численности населения стран мира используем информацию международной организации The United Nations (The United States, New York City) [5].

3. Выявим и исключим выбросы данных Пл в выборке. Для этого применим метод межквартильного размаха и сводки пяти чисел.

Получим: выборочный минимум - 0,06; первый квартиль - 0,89;

---

медиана - 4,77; третий квартиль - 21,63; выборочный максимум - 544,79. Межквартильный размах, умноженный на коэффициент  $k$  (1,5) – 31,11; нижняя внутренняя граница = -30,22, данных ниже в выборке нет. Верхняя внутренняя граница = 52,74; 13 значений данных выше границы.

Также для выявления выбросов используем метод диаграммы.

4. Удалим выявленные выбросы. Это данные 10 стран, определенных по совокупности методов. Продолжим далее с выборкой 149 стран.

5. Проранжируем страны таким образом, что страны с наименьшей Пл получит ранг 1, а с наибольшей – 149.

Продолжим далее с выборкой 149 стран.

6. Проверку гипотезу  $H_0$  о существовании значимой обратной связи между значениями  $Y$  (Пл) и уровнем фактора  $A$  (КБ) начнем с корреляционного анализа связи Пл и Глобального индекса GCI международной организации в сфере информационно-коммуникационных технологий The International Telecommunication Union (ITU) (Switzerland, Geneva) [6].

GCI рассчитывается на основе ответов 193 государств – членов ООН на 82 вопроса, объединённых по 5 направлениям мер, направленных на обеспечение КБ: правовых, технических, организационных, развития потенциала, сотрудничества.

По расчету, коэффициент корреляции показателей Пл и GCI составляет 0,58, то есть, по результатам корреляционного анализа выявлено, что низкому уровню Пл со средней силой соответствует низкий уровень КБ, а высокому уровню утечек ПД - более высокая оценка индекса КБ. Следовательно, для связи Пл и GCI гипотеза  $H_0$  о существовании значимой обратной связи между значениями  $Y$  (страновые утечки ПД) и уровнем фактора  $A$  (КБ) опровергнута.

---

7. Далее проверим верность гипотезы  $H_0$  для связи между значениями Пл и Национальным индексом КБ (NCSI), рассчитываемым некоммерческим фондом The E-Governance Academy (Estonia, Tallinn) [7].

По утверждению разработчика NCSI – это глобальный индекс, который измеряет готовность стран к предотвращению кибернарушений и управлению киберинцидентами. Индекс построен на основе 49 показателей, распределенных по 3 категориям (общие показатели, базовые показатели КБ, показатели управления инцидентами и кризисными ситуациями), каждая из которых описывает 4 возможности. В рейтинг включены данные по 176 странам.

Рассчитанный коэффициент корреляции Пл и NCSI составил 0,67, то есть связь показателей имеет среднюю положительную величину. Следовательно, и для связи Пл и NCSI гипотеза  $H_0$  о существовании значимой обратной связи между значениями  $Y$  (Пл) и уровнем фактора  $A$  (КБ) опровергнута и подтверждена альтернативная гипотеза  $H_1$  об отсутствии такой связи. Коэффициенты корреляции между Пл и индексами КБ приведены в таблице № 1.

Таблица № 1

Коэффициенты корреляции связи Пл с показателями КБ

Показатель	Коэффициент корреляции
Глобальный индекс КБ GCI	0,58
Национальный индекс КБ NCSI	0,67

Таким образом, согласно расчету развитые страны с высокими оценками выполнения мер по обеспечению КБ одновременно являются и лидерами по Пл. Такое явление можно оценить, как парадокс, который требует отдельного исследования. Возможным направлением является работа

по совершенствованию расчетов индексов КБ, направленную на обеспечение соответствия между расчетным и фактическим уровнями КБ.

8. Далее проверим гипотезу  $H_{01}$  о существовании взаимодействия фактора А (КБ) и фактора В (не установлен), влияющего на значения  $Y$ . Для определения фактора В анализируем причины, влияющие на  $Y$ . Поиск осуществляем среди мотивов киберпреступников – взломщиков аккаунтов.

Например, А.У. Менциев и Х.С. Чебиева разделяют мотивы на 2 группы. К первой относятся мотивы киберпреступников, имеющих серьезные намерения на организацию значительных сбоев, нанесение финансовых ущербов, мошенничество и кражу данных. Во вторую группу входят люди, имеющие интерес к самому процессу взлома, бросающих вызов системе, для некоторых из них – это своеобразное хобби [8].

Также в качестве кибернарушителей могут быть не отдельные лица, а целые организации и даже правительства стран. Такое наблюдается в случае активизации киберпротивостояния на фоне геополитических конфликтов. Так, И. В. Лапшина и Е. А. Першонкова показывают двойственность кибервойны как военного конфликта и как сетевой войны, имеющей социальную направленность. В последнем случае противники стремятся захватить контроль и над домашними компьютерами, воздействуя через сеть Интернет, проникая в электронную почту, организуя фишинг [9].

9. После анализа мотивов остановимся на факторе высокой заинтересованности киберпреступников в получении финансовой выгоды, что объясняет их целенаправленность на граждан развитых стран. Выберем в качестве показателя этого фактора ВВП на душу населения.

Проверим корреляцию  $Pl$  с ВВП на душу населения (ППС, текущие международные доллары). В качестве источника возьмем данные международного финансового учреждения The World Bank (The United States,

---

Washington, D.C.) [10]. Получившийся в результате корреляционного анализа коэффициент, равный 0,66, характеризует связь между Пл и ВВП на душу населения как прямую средней силы. Следовательно, странам с наименьшей Пл соответствуют меньшие значения показателя ВВП на душу населения, а странам с большей Пл – более высокий уровень благосостояния населения.

10. Рассчитаем реальные значения фактора А, очищенные от влияния фактора В (ВВП на душу населения). Расчет произведем для Глобального индекса КБ (GCI). Для этого произведем с показателями следующие действия.

А. Выразим взаимодействие факторов А (КБ) и В (ВВП на душу населения) через их произведение.

В. Чтобы элиминировать влияние фактора В (ВВП на душу населения) на фактор А (КБ) примем базовое значение фактора В равным 1:  $B_0 = 1,00$ .

С. Уравнение  $Y$  примет вид формулы (1).

$$Y = A_i * B_i = A_i * (B_0 * k_{B_i}) = (A_i * k_{B_i}) * B_0 = A_i * k_{B_i}, \quad (1)$$

где  $Y$  – Пл;

$A_i$  –  $i$ -е реальное значение А в наборе данных, очищенное влияния фактора В (ВВП на душу населения);

$B_i$  –  $i$ -е значение В в наборе данных;

$k_{B_i}$  – коэффициент  $B_i$ -го значения, определяемый как отношение значений  $B_i$  к  $B_0$ .

Д. Тогда А (КБ) из рейтинга ITU определяется по формуле (2).

$$A_{ITU} = A_i * k_{B_i}, \text{ откуда } A_i = \frac{A_{ITU}}{k_{B_i}}, \quad (2)$$

где  $A_{ITU}$  – значения фактора А (КБ), указанные в рейтинге ITU.

Е. Рассчитаем реальные значения GCI, NCSI для всех 149 стран в выборке по формуле (2).

Ф. Проведем корреляционный анализ связи между Пл и пересчитанными GCI, NCSI.

Г. Занесем полученные коэффициенты корреляции в таблицу № 2.

По результатам элиминации фактора В (ВВП на душу населения) и корреляционного анализа связи Пл и пересчитанных показателей КБ можно сделать вывод, что гипотеза  $H_{01}$  о существовании взаимодействия фактора А (КБ) и фактора В, влияющего на значения  $Y$  (Пл), подтверждена.

Таблица № 2

Коэффициенты корреляции связи Пл с пересчитанными показателями КБ

Показатель	Коэффициент корреляции	
	прежний	пересчитанный
Глобальный индекс КБ GCI	0,58	-0,43
Национальный индекс КБ NCSI	0,67	-0,48

В данном исследовании выдвинуты гипотезы о существовании значимой обратной связи между значениями  $Y$  (страновые утечки ПД) и уровнем фактора А (КБ), о существовании взаимодействия фактора А (КБ) и иных факторов, влияющего на значения  $Y$  (страновые утечки ПД).

В результате корреляционного анализа подтверждено наличие фактора В (ВВП на душу населения), взаимодействующего с фактором А (КБ) и оказывающего влияние на  $Y$  (Пл). После элиминации влияния фактора В подтверждена обратная связь  $Y$  с показателями КБ (GCI; NCSI).

В ходе исследования было высказано предположение наличия проблем с полнотой факторов, учтенных в показателях КБ и достоверностью данных, получаемых от правительств стран, что требует дальнейших исследований.



## Литература

1. Утечки информации в мире, 2022-2023 годы // INFOWATCH. URL: [infowatch.ru/analytics/analitika/issledovaniye-utechek-informatsii-v-mire-za-posledniye-dva-goda](http://infowatch.ru/analytics/analitika/issledovaniye-utechek-informatsii-v-mire-za-posledniye-dva-goda).
  2. Баландин Ю.А. Кибербезопасность и информационная безопасность. Демаркация правовых категорий // Правовая политика и правовая жизнь. 2023. №3. С. 260-270.
  3. Explainers // IBM URL: [ibm.com/topics](http://ibm.com/topics).
  4. Data breach statistics by country: 2023vs2022 // Surfshark. URL: [surfshark.com/research/study/data-breach-recap-2023](http://surfshark.com/research/study/data-breach-recap-2023).
  5. World Population Prospects 2024 // United Nations. URL: [population.un.org/wpp/Download/SpecialAggregates/EconomicTrading/](http://population.un.org/wpp/Download/SpecialAggregates/EconomicTrading/).
  6. Global Cybersecurity Index 2020: Measuring commitment to cybersecurity // ITU URL: [itu.int/dms\\_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf](http://itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf).
  7. Archived data from 01.09.2023 // NCSI: national cyber security index. URL: [ncsi.ega.ee/ncsi-index/?archive=1](http://ncsi.ega.ee/ncsi-index/?archive=1).
  8. Менциев А.У., Чебиева Х.С. Современные угрозы безопасности в сети Интернет и контрмеры (обзор) // Инженерный вестник Дона, 2019, №3. URL: [ivdon.ru/ru/magazine/archive/N3y2019/5859](http://ivdon.ru/ru/magazine/archive/N3y2019/5859).
  9. Лапшина И.В., Першонкова Е.А. Рефлексивно управляемые кибервойны современности с позиции когнитивного моделирования // Инженерный вестник Дона, 2021, №9. URL: [ivdon.ru/ru/magazine/archive/n9y2021/7187](http://ivdon.ru/ru/magazine/archive/n9y2021/7187).
  10. GDP per capita, PPP (current international \$) // World Bank Group URL: [data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?end=2023&start=2023&view=bar&year=2023](http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?end=2023&start=2023&view=bar&year=2023).
-



## References

1. Utechki informacii v mire, 2022-2023 gody [Information leaks in the world, 2022-2023] URL: [infowatch.ru/analytics/analitika/issledovaniye-utechek-informatsii-v-mire-za-posledniye-dva-goda](http://infowatch.ru/analytics/analitika/issledovaniye-utechek-informatsii-v-mire-za-posledniye-dva-goda).
2. Balandin Ju.A. Pravovaja politika i pravovaja zhizn'. 2023. №3. pp. 260-270.
3. Explainers URL: [ibm.com/topics](https://ibm.com/topics).
4. Data breach statistics by country: 2023vs2022. URL: [surfshark.com/research/study/data-breach-recap-2023](https://surfshark.com/research/study/data-breach-recap-2023).
5. World Population Prospects 2024. URL: [population.un.org/wpp/Download/SpecialAggregates/EconomicTrading/](https://population.un.org/wpp/Download/SpecialAggregates/EconomicTrading/).
6. Global Cybersecurity Index 2020. Measuring commitment to cybersecurity URL: [itu.int/dms\\_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf](https://itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf).
7. Archived data from 01.09.2023. URL: [ncsi.ega.ee/ncsi-index/?archive=1](https://ncsi.ega.ee/ncsi-index/?archive=1).
8. Menciev A.U. Chebieva H.S. Inzhenernyj vestnik Dona, 2019, №3. URL: [ivdon.ru/ru/magazine/archive/N3y2019/5859](https://ivdon.ru/ru/magazine/archive/N3y2019/5859).
9. Lapshina I.V., Pershonkova E.A. Inzhenernyj vestnik Dona, 2021, №9. URL: [ivdon.ru/ru/magazine/archive/n9y2021/7187](https://ivdon.ru/ru/magazine/archive/n9y2021/7187).
10. GDP per capita, PPP (current international \$). URL: [data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?end=2023&start=2023&view=bar&year=2023](https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?end=2023&start=2023&view=bar&year=2023).

**Дата поступления: 27.07.2024**

**Дата публикации: 5.09.2024**