УДК 519.2

ПАРАДОКС ДНЕЙ РОЖДЕНИЯ В КРИПТОГРАФИИ

Наталия Владимировна Пчелинцева[1]
старший преподаватель
natas79@mail.ru
Кирилл Олегович Самохин[1]
студент
mansuha@bk.ru
Ольга Сергеевна Картечина[2]
студент
kartechnatali@mail.ru
[1]Мичуринский государственный аграрный университет
г. Мичуринск, Россия
[2]Российский университет транспорта (МИИТ)
г. Москва, Россия

Аннотация. В статье рассматривается парадокс дней рождения и его использование в криптографии для разработки атак на хеш-функции, рассматривается суть метода атак «дней рождения», а также другие явления с аналогичной структурой.

Ключевые слова: криптография, парадокс дней рождения, цифровые подписи, хэш-функции, шифрование.

Парадокс дней рождения — утверждение, состоящее в том, что в группе, состоящей из 23 или более человек, вероятность совпадения дней рождения (число и месяц) хотя бы у двух людей превышает 50 %. Например, если в группе 23 человека или более, то более вероятно то, что у какой-то пары людей дни рождения придутся на один день, чем то, что у каждого будет свой неповторимый день рождения. Впервые эта задача была рассмотрена Рихардом Мизесом в 1939 году.

Для 57 и более человек вероятность такого совпадения превышает 99%, хотя 100% она достигает, согласно принципу Дирихле, только тогда, когда в группе не менее 367 человек (ровно на 1 больше, чем число дней в високосном году; с учётом високосных лет).

Такое утверждение может показаться неочевидным, так как вероятность совпадения дней рождения двух человек с любым днём в году 1/365=0,27%, умноженная на число человек в группе (23), даёт лишь 1/365*23=6,3%. Это рассуждение неверно, так как число возможных пар 23* 22/2=253 значительно превышает число человек в группе (253>23). Таким образом, утверждение не является парадоксом в строгом научном смысле: логического противоречия в нём нет, а парадокс заключается лишь в различиях между интуитивным восприятием ситуации человеком и результатами математического расчёта.

Мы можем обобщить парадокс дней рождения, чтобы взглянуть на другие явления с аналогичной структурой [1].

Вероятность того, что у двух людей будет один и тот же PIN-код на их банковской карте, составляет 1:10000, или 0,01%. Однако потребовалась бы всего группа из 119 человек, чтобы иметь шансы в пользу двух человек, имеющих один и тот же PIN-код.

Конечно, эти числа предполагают случайную выборку, равномерное распределение дней рождения и PIN-кодов. На самом деле пик дней рождения приходится на определенное время года, и люди с большей вероятностью выберут определенные числа для своего PIN-кода, чем другие. Но отсутствие равномерного распределения фактически уменьшает размер нужной группы.

Если мы уменьшим вероятность совпадения, то размер группы, необходимый для получения равной вероятности совпадения, очевидно, увеличится. Однако она увеличивается гораздо медленнее, чем обратная вероятность [2].

Например, с вероятностью 1:10000 минимальный размер группы составляет 119 человек. Для совпадения, вероятность которого в 10 раз меньше, минимальная группа составляет 373, или всего в 3,15 раза больше. Следовательно, даже при невероятно малых вероятностях размер группы не становится особенно большим. При коэффициенте один к миллиону требуемая группа составляет всего 1178 человек.

Космический мусор. Чем больше космического мусора, тем выше вероятность столкновений. И каждое столкновение увеличивает количество кусков космического мусора. Эта петля положительной обратной связи, если она превышает скорость, с которой объекты попадают в атмосферу и сгорают, может привести к так называемому синдрому Кесслера. Это цепная реакция, в ходе которой столкновения становятся все более частыми, разбрызгивая все больше и больше мусора, пока размещение спутника на низкой околоземной орбите не станет слишком опасным, чтобы быть осуществимым.

Вероятность того, что два конкретных орбитальных объекта столкнутся друг с другом в течение года, почти бесконечно мала. Однако, учитывая, что над нашими головами проносится около 5500 спутников и примерно 900 000 объектов размером более 1 см, столкновения происходят чаще, чем можно ожидать.

В 2009 году два спутника — 16-летний несуществующий российский военный спутник и все еще активный спутник связи Iridium - столкнулись с относительной скоростью почти 12 км/с. Оба спутника разлетелись на облака осколков, более 1000 осколков размером больше грейпфрута.

Доказательства ДНК. За последние сорок лет ДНК-улики произвели революцию в области судебных расследований. Занимаясь своими повседневными делами, мы оставляем за собой след генетического материала, в

основном через клетки кожи и волосы. Правительства составляют огромные базы данных «профилей» ДНК, регистрируя ряд некоррелированных генетических маркеров.

Для некоторых систем вероятность совпадения двух людей по всем зарегистрированным генетическим маркерам оценивается в один к одному триллиону (исключая идентичных близнецов).

Исходя из предыдущих примеров, крошечная вероятность может превратиться во что-то осязаемое, когда у вас имеется достаточно большая группа людей.

В стране с населением 328 миллионов человек коэффициент совпадения один к триллиону преобразуется в 1 к 3000 шансов на то, что у вас есть генетический профиль «близнец» где-то там. Например, в 2019 году в США было совершено 16 тысяч убийств. Это означает, что, вероятно, происходит около 5 убийств в год, для которых ДНК преступника идеально совпадает с ДНК другого американца (опять же, исключая идентичных близнецов). Даже с учетом невероятно низкой вероятности, сила парадокса дней рождения означает, что вы не должны выносить обвинительный приговор только на основании доказательств ДНК, а также необходимо принимать во внимание другие косвенные доказательства.

Стоит также учитывать, что системы профилирования ДНК значительно улучшились за последние тридцать лет. Ранее при применении этой технологии часто указывались вероятности 1 на миллиард. Это дало бы около 5000 убийств с двусмысленностью ДНК.

Атака на день рождения. Парадокс дней рождения может быть использован в криптографической атаке на цифровые подписи. Цифровые подписи основаны на так называемой хэш-функции f (x), которая преобразует сообщение или документ в очень большое число (хэш-значение). Затем этот номер объединяется с секретным ключом подписывающего лица для создания подписи. Кто-то, читающий документ, может затем «расшифровать» подпись,

используя открытый ключ подписавшего, и это докажет, что подписавший подписал документ цифровой подписью [3].

Эти подписи могут быть использованы для проверки подлинности документа. Прочитав любую статью, вы используете цифровую подпись прямо сейчас, по протоколу HTTPS. Безопасность зависит от сложности поиска другого документа с тем же хэш-значением, что и подписанный оригинал.

Однако парадокс дней рождения позволяет нам потенциально злоупотреблять этой системой, атакуя эту хэш-функцию.

Допустим, А. - это орган, который подписывает контракты в цифровом виде. Мы хотим обманом заставить А. подписать мошеннический контракт, не зная об этом, чтобы позже мы могли предположить, что он его одобрил. Что нам нужно найти, так это два контракта, один законный и один мошеннический, которые при передаче через f(x) выдают одно и то же хэш - значение.

Для каждого контракта мы можем определить множество способов незаметного изменения его, не изменяя его значения. Например, вы можете добавить различное количество пробелов в конце каждой строки, слегка изменить пиксели в логотипе или внести небольшие изменения в форматирование. В сочетании это дает нам миллионы технически разных, но семантически идентичных документов, которые, по мнению А., все получили бы печать одобрения. Это также дает нам миллионы вариантов поддельного документа. Если мы найдем пару документов, один законный, один мошеннический, которые выдают один и тот же хэш, тогда мы можем передать законный А. для подписания, а затем использовать эту подпись, чтобы «доказать» подлинность мошеннического контракта [4, 5].

Благодаря парадоксу дней рождения вероятность по крайней мере одного столкновения хэш - значений между одним из законных и одним из поддельных документов намного выше, чем можно было бы ожидать, учитывая огромный диапазон хэш-функции τ[1, 3, 6]. Фактически, количество документов, которые вам нужно создать, составляет примерно квадратный корень из числа

возможных выходных данных хэш-функции. Это улучшается за счет того факта, что ни одна хэш-функция не распределена идеально равномерно, что привело к тому, что многие популярные алгоритмы хеширования стали небезопасными [5].

Сходный математический аппарат используется для оценки размера популяции рыб, обитающих в озёрах. Метод называется «capture-recapture» («поймать — поймать снова»). Действительно, если каждую пойманную рыбу помечать и отпускать, то вероятность поймать помеченную рыбу будет расти нелинейно с ростом количества попыток. Размер популяции грубо может быть оценён как квадрат числа попыток, совершаемых до вылавливания первой помеченной рыбы.

Список литературы:

1. Фомичёв В.М., Мельников Д.А. Криптографические методы защиты информации в 2 ч. Часть 1. Математические аспекты: учебник для академического бакалавриат: под редакцией В. М. Фомичёва. - Москва: Издательство Юрайт, 2019
2. Умарзода С.У. Этапы развития криптографии и стеганографии // В сборнике: Права человека в современном мире: концепции, реальность и перспективы. Материалы международной научно-практический конференции, посвящённой Дню прав человека и международному дню борьбы с коррупцией. Душанбе, 2022. С. 404-414.
3. Пчелинцева Н.В. Методические аспекты количественной оценки риска в аграрной сфере производства // Наука и Образование. 2019. № 3. С. 37.
4. Дегтярева А.А., Пчелинцева Н.В, Макова Н.Е. Математические основы криптологии // Наука и Образование. 2020. Т. 3. № 2. С. 46.
5. Иванов С.Г., Доротскар З. Профессиональный соперник криптографии (ПСК): модель разработки игр для изучения криптографии // Международная конференция по мягким вычислениям и измерениям. 2021. Т. 1. С. 312-315.

6. Заболотникова М.А., Картечина О.С., Пчелинцева Н.В Сравнительный анализ хэш-функций // Наука и Образование. 2020. Т. 3. № 2. С. 48.

UDC 519.2

THE BIRTHDAY PARADOX IN CRYPTOGRAPHY

Natalia V. Pchelintseva[1]
Senior lecturer
natas79@mail.ru
Kirill O. Samokhin[1]
student
mansuha@bk.ru
Olga S. Kartechina[2]
student
kartechnatali@mail.ru
[1]Michurinsk State Agrarian University
Michurinsk, Russia
[2]Russian University of Transport (MIIT)
Moscow, Russia

Annotation. The article discusses the paradox of birthdays and its use in cryptography for the development of attacks on hash functions, examines the essence of the method of attacks of "birthdays", as well as other phenomena with a similar structure.

Key words: cryptography, birthday paradox, digital signatures, hash functions, encryption.

Статья поступила в редакцию 29.03.2022; одобрена после рецензирования 11.04.2022; принята к публикации 12.05.2022.

The article was submitted 29.03.2022; approved after reviewing 11.04.2022; accepted for publication 12.05.2022.