УДК 004.056.5 **DOI:10.30837/rt.2019.3.198.11**

Н.А. ПОЛУЯНЕНКО, канд. техн. наук, А.А. КУЗНЕЦОВ, д-р техн. наук

МОДЕЛИРОВАНИЕ АТАКИ ДВОЙНОЙ ТРАТЫ НА ПРОТОКОЛ КОНСЕНСУСА «PROOF OF WORK»

1. Введение

Наиболее известной атакой на протоколы консенсуса блокчейн-систем является так называемая атака двойной траты (двойного расходования, англ. Double-spending), когда нечестный участник децентрализованной системы осуществляет повторное отчуждение (продажу) одних и тех же цифровых активов (единиц криптовалюты, токенов, монет и пр.), т.е. реализует несколько незаконных платежей из одного и того же стартового состояния [1, 2]. Если между заключением сделки и оформлением передачи права собственности проходит значительный промежуток времени, тогда продавец может попытаться продать один и тот же товар несколько раз разным покупателям, получая несколько раз оплату за один и тот же актив. Наиболее актуальной задача предотвращения двойной траты становится в системах электронных платежей. Цифровые активы легко копируются, и нечестный участник может передавать их копии большому количеству клиентов. Каждый получатель может убедиться, что полученный актив полностью соответствует заявленным характеристикам, однако не будет уверен, что такой же копией не расплатились с другим участником системы.

В традиционных (централизованных) системах задачу предотвращения двойной траты решают применением административных мер, когда централизованный (которому все подчинены) узел обеспечивает контроль допустимости той или иной операции. За предотвращение двойной траты в децентрализованной распределенной системе отвечают протоколы принятия консенсуса относительно того, какую транзакцию считать истинной [1 – 3]. Этот механизм позволяет (в идеальном случае) игнорировать попытки двойного расходования одних и тех же цифровых активов.

Первым и наиболее изученным протоколом консенсуса децентрализованных систем является алгоритм «Proof of work» [3, 4]. В его основе лежит решение сложной вычислительной задачи (как правило, поиск прообраза функции криптографического хеширования). И только тот, кто первым решит эту задачу (найдет подходящий прообраз), получит право внести изменение в состояние системы [4]. Фактически это означает возможность осуществлять транзакции с отчуждением (продажей, оплатой и пр.) цифровых активов. Таким образом, задача предотвращения двойной траты состоит в исключении (или, по крайней мере, снижении вероятности) возможного формирования прообраза одним и тем же участником системы. На практике это достигается вовлечением огромного числа участников с соответствующим распределением их вычислительных возможностей по поиску прообразов криптографической функции хеширования. Дополнительно каждый участник вправе передавать права на свои активы только после некоторого числа сформированных прообразов, кратно снижая тем самым вероятность двойной траты.

Первые результаты по оценке вероятности двойной траты в децентрализованной системе Биткоин были опубликованы в оригинальной статье Сатоши Накамото [4], а также в работе Мени Розенфельда [5]. Это самые популярные и цитируемые работы в данной области. Существуют также другие публикации, которые для разных случаев уточняют и дополняют результаты, полученные С. Накамото и М. Розенфельдом:

– результаты Карлоса Пинзон и Камило Роча [6], строящие модели атак двойной траты на основе не только хешрейта (вычислительных возможностей) злоумышленника и честной сети, но учитывающие также влияние временных параметров. Уравнения, управляющие этими моделями, используют распределение вероятностей Эрланга, в отличие от работы С. Накамото, который использует распределение вероятностей Пуассона, и работы

М. Розенфельда, который использует отрицательное биномиальное распределение вероятностей;

– результаты Ковальчук [7], которые обобщают и частично развивают известные оценки, также учитывающие время подтверждения транзакции;

– работу Аззолини [8], в которой используется вероятностное логическое программирование. Как утверждается, данный метод позволяет учитывать переменную во времени скорость хеширования и переменную сложность алгоритма «Proof of work»;

– результаты Кевин Ляо, представленные в работе [9] и рассматривающие китовую атаку, в которой злоумышленник из числа меньшинства увеличивает свои шансы на успешное проведение атаки двойной траты, стимулируя майнеров подорвать согласованный протокол и вступать в сговор посредством китовых транзакций или транзакций, несущих аномально большие сборы.

Следует отметить, что известные оценки получены в результате некоторых упрощений и допущений, т.е. используемые модели, как правило, дают приближенные значения, и основная критика этих оценок состоит в их нереалистичности, оторванности от реальных процессов, протекающих в децентрализованных системах. В частности, островными неточностями и ложными допущениями в работах С. Накамото и М. Розенфельда являются:

– вероятности сформировать блок честной сетью и злоумышленником в сумме должны быть равны единицы. Однако приведенные выражения не дают ответа, какой будет результат при независимых величинах этих вероятностей [10];

– не принимается во внимание экономическая возможность по формированию блоков злоумышленником, а также экономическая целесообразность. Ресурсы злоумышленника по поддержанию гонки между злоумышленником и честной сетью считаются безграничными, что не может соответствовать действительности [11];

– предполагается, что вероятность успеха сформировать блок не меняется во время эксперимента, хотя, в действительности, майнеры могут изменить свои вероятности поиска нужного прообраза и формирования блока, увеличивая или уменьшая свои вычислительные ресурсы [11];

– в работе М. Розенфельда теорема про вероятность успеха злоумышленником приведена без доказательства и получена с допущением о времени распространения блока в сети равным нулю, в [12] упомянуто о том, что нужно учитывать время синхронизации сети;

– допущение о формировании блоков в соответствии со средним временем ожидания блока, сделанное в работе С. Накамото, ошибочно [13].

К сожалению, существует не так много работ, в которых проводится попытка экспериментально подтвердить или опровергнуть полученные теоретические расчеты, т.е. эмпирическим путем обосновать адекватность выбранной математической модели. К таким работам можно отнести [8, 11].

Во всех упомянутых работах используется модель разорения игрока, проверяемая методами Монте-Карло. На основе данной модели и выводится формула для расчета вероятности успешного проведения атаки двойной траты.

Цель данной работы – критический анализ известных аналитических оценок вероятности успешной реализации атаки двойной траты на протокол консенсуса «Proof of work». В частности, мы рассматриваем «задачу о разорении игрока», лежащую в основе моделей С. Накамото и М. Розенфельда, и показываем, что базовые предположения о вероятностном пространстве (множество элементарных исходов и вероятности их наступления) не соответствуют реальным процессам, протекающим при установлении консенсуса «Proof of work» в блокчейн-системе. Далее, для теоретической оценки вероятности успешной атаки двойной траты мы предлагаем использовать модель «независимых игроков», которая, на наш взгляд, устраняет основные неточности и несоответствия. Эмпирическим путем мы показываем сходимость результатов теоретических расчетов с данными экспериментов по имитации «гонки» между честными игроками и злоумышленниками. Наиболее интересным, на наш взгляд,

является сопоставление результатов теоретических расчетов, полученных применением различных моделей, и эмпирических результатов, полученных имитацией «гонки».

2. «Задача о разорении игрока» применительно к атаке двойной траты

Рассмотрим «задачу о разорении игрока», а точнее ее небольшую модификацию, на которую ссылается С. Накамото, цитируя известный учебник Феллера 1968 г. [14], или М. Розенфельд, моделируя процесс гонки как процесс эквивалентный цепочке Маркова с дискретным временем, где каждый шаг определяется как поиск блока кем-либо.

Приведем сначала выдержку из раздела 11 работы С. Накамото [4], в которой излагаются рассуждения в отношении моделирования атаки двойной траты:

«Гонку между честными участниками и нападающим можно представить как биномиальное случайное блуждание. Успешное событие, когда «честная» цепь увеличивается на один блок, приводит к увеличению отрыва на единицу, увеличивая свое преимущество на $+1$, а неуспешное, когда очередной блок создает злоумышленник, – к его сокращению на один блок, уменьшая разрыв на -1. Вероятность атакующего наверстать разницу в несколько блоков такая же, как и в задаче о «разорении игрока». Представим, что игрок имеет неограниченный кредит, начинает с некоторым дефицитом и у него есть бесконечно много попыток, чтобы отыграться».

Далее приведем выдержку из раздела 3 работы М. Розенфельда [5]:

«Обозначим через $z = n - m$ количество блоков, в которых честная сеть имеет преимущество перед атакующим. Всякий раз, когда блок найден, значение z изменяется; если этот блок был обнаружен честной сетью, z увеличивается на 1, а если этот блок был обнаружен атакующим, z уменьшается на 1. Формально это цепочка Маркова с непрерывным временем и скоростью p/T_0 для продвижения вверх на шаг, и скорость q/T_0 для движения вниз на шаг».

Как видим, в этих работах используются модель, в которой в каждом испытании выигрывает злоумышленник (формируя очередной блок) или злоумышленник проигрывает и при этом считается, что выигрывает честная сеть (формируя очередной блок). Однако в статьях не приводится какого-либо обоснования выбранной модели. Авторы допускают, что если блок не сформировал злоумышленник, то, в таком случае, блок обязательно формирует честная сеть, никак не обосновывая это допущение.

Действительно, в определении задачи о разорении игрока используется вероятностное пространство с двумя элементарными событиями: «выиграл первый игрок»; «выиграл второй игрок». При моделировании атаки двойной траты С. Накамото и М. Розенфельд интерпретируют элементарные исходы этой задачи как «блок сформирован честной сетью» (по традиции вероятность такого исхода обозначается p) и «блок сформирован атакующим» (с вероятностью q), причем $p = 1 - q$. Однако в реальных блокчейн-системах вероятность формирования блока (нахождения прообраза функции хеширования) определяется исключительно хешрейтом (вычислительными возможностями) каждого участника, т.е. условие $p = 1 - q$ не обязано выполняться. Например, при хешрейте участников, превышающем сложность поиска прообраза за определенный интервал времени, каждый участник гарантированно найдет прообраз, т.е. сформирует блок и, в этом случае, $p = 1$ и $q = 1$. В реальных системах сложность поиска прообраза корректируется исходя из вычислительных возможностей участников, причем так, чтобы прообраз был найден за определенный временной интервал (например, в криптовалюте биткоин это 10 мин.). Если предположить, что такая корректировка выполняется над двумя игроками: «честная сеть» и «атакующий», а p и q – соответствующие вероятности формирования блока за определенный временной интервал, тогда предположение $p = 1 - q$ оправданно. Однако в реальной ситуации злоумышленник атакует систему, не оглашая своих вычислительных возможностей и, вероятнее всего, скрывая сам факт предполагаемой атаки, т.е. предположение $p = 1 - q$ не имеет оснований.

Если оставить введенные обозначения (вероятности p и q) и отказаться от обязательного выполнения условия $p = 1 - q$, тогда в результате каждой попытки (или серии попыток в течение заданного интервала времени) пространство элементарных исходов содержит такие элементарные события:

- «блок сформирован честной сетью и атакующий не сформировал блок» с вероятностью $p(1-q)$;
- «блок не сформирован честной сетью и атакующий сформировал блок» с вероятностью $(1-p)q$;
- «блок не сформирован честной сетью и атакующий не сформировал блок» с вероятностью $(1-p)(1-q)$;
- «блок сформирован честной сетью и атакующий сформировал блок» pq.

Множество всех элементарных исходов составляет полную группу событий:

$$p(1-q) + (1-p)q + (1-p)(1-q) + pq = 1.$$

Эта модель с четырьмя элементарными исходами (будем называть ее в дальнейшем «модель с независимыми игроками») описывает реальный вероятностный процесс в блокчейн-системе при установлении консенсуса на основе алгоритма «Proof of work».

3. Сравнение вероятностных событий в двух исследуемых моделях

В модели независимых игроков формирование очередного блока у злоумышленника и честной сети происходит независимо друг от друга, вероятности поиска прообраза хеш-функции (для формирования блока) определяются их хешрейтами (вычислительными возможностями). Для сравнения с результатами, полученными в работах С. Накамото и М. Розенфельдом (для модели разорения игрока), будем использовать общепринятые упрощения:

– время распространения блока по сети пренебрежимо мало, т.е. обмен информацией между узлами происходит практически мгновенно (время синхронизации равно нулю);

– хешрейт злоумышленника, хешрейт честной сети и сложность майнинга не меняются со временем на протяжении всей гонки;

– возможности злоумышленника по поддержанию состояния гонки достаточно велики, но не бесконечны;

– кроме злоумышленника все остальные пользователи сети действуют строго в соответствии с правилами протокола блокчейн-сети;

– победой злоумышленника будем считать формирование необходимого количества блоков подтверждения раньше или одновременно (считается, что один блок злоумышленник сформировал заранее) или в противном случае – последующего формирования цепочки блоков равной длины с честной сетью.

В задаче двойной траты злоумышленник выигрывает, если сформирует равное с честной сетью количество блоков, при условии, что честная сеть уже сформировала N блоков. Здесь используем ту же формулировку, что и в работе М. Розенфельда [5], предполагая, что один блок был предварительно добыт атакующим до начала атаки и, следовательно, общая длина сформированной злоумышленником цепочки будет на единицу больше, что является достаточным условием для принятия ее честной сетью как основной блокчейн.

Если предполагать, что ресурсы у злоумышленника конечны или выигрыш злоумышленником не покроет его финансовых затрат на поддержание дальнейшей гонки, то логично предположить об ограничении на формирование максимального количества блоков в состязании [11]. Предположим, что злоумышленник отказывается от продолжения гонки в случае, если честная сеть сформировала $N + n_{\max}$ блоков. Все состояния, в которых злоумышленник не выиграл, будут для него проигрышными.

Необходимо обратить внимание, на два момента в модели разорения игрока:

1) злоумышленник не может выиграть гонку раньше, чем за $2 \cdot N$ попыток (необходимо не менее N попыток для формирования N блоков честной сетью и столько же попыток для формирования N блоков злоумышленником);

2) злоумышленник может выиграть при нечетном количестве попыток, только если он опередил честную сеть до того, как она сформировала N блоков (вероятность чего значительно меньше при меньших мощностях майнинга злоумышленником).

В отличие от модели разорения игрока, в модели независимых игроков злоумышленник может выиграть, начиная с N попыток и так как события формирования блоков обоими участниками независимы и нет зависимости вероятности выигрыша от четности или нечетности текущей попытки.

Рассмотрим пример расчета вероятности наступления какого-либо события для двух рассмотренных моделей. Для определенности положим вероятность формирования блока злоумышленником за каждую попытку $q = 0,3$ (вероятность не сформировать блок будет $(1-q) = 0,7$). Для согласования с моделью разорения игрока положим $p = 0,7$ (вероятность не сформировать блок честной сетью будет $(1-p) = 0,3$). Необходимое количество подтверждений $N = 1$. Ограничение на формирование максимального количества блоков в состязании $N + n_{\max} = 2$ попытки.

Проанализируем вероятности различных исходов для различных моделей.

Рассмотрим модель разорения игрока:

1) первая попытка (два возможных исхода):

– формирование блока честной сетью, злоумышленник отстает на один блок, гонка продолжается, вероятность наступления такого события равна $p = 0,7$;

– формирование блока злоумышленником, честная сеть отстает на один блок, победа злоумышленника[1], вероятность наступления события равна $q = 0,3$;

2) вторая попытка (четыре возможных исхода, рассматриваем только случай формирования блока честной сетью в первой попытке, т.е. когда гонка продолжается):

– и в первой, и во второй попытке сформирован блок честной сетью, злоумышленник проиграл в гонке, гонка завершена, вероятность наступления события $p \cdot p = 0,49$;

– в первой попытке сформирован блок честной сетью, но во второй попытке блок сформирован злоумышленником, злоумышленник выиграл, гонка завершена, вероятность наступления события $p \cdot p = 0,21$.

Таким образом, в модели разорения игрока злоумышленнику удастся победить с вероятностью $0,3 + 0,21 = 0,51$.

Для модели независимых игроков:

1) первая попытка (четыре возможных исхода):

– блок сформирован честной сетью и атакующий не сформировал блок, злоумышленник отстает на один блок, гонка продолжается, вероятность наступления события $p \cdot (1-q) = 0,49$;

– блок не сформирован честной сетью и атакующий сформировал блок, победа злоумышленника[1], вероятность наступления события $(1-p) \cdot q = 0,09$;

– блок не сформирован честной сетью и атакующий не сформировал блок, гонка продолжается, вероятность наступления события $(1-p) \cdot (1-q) = 0,21$;

– блок сформирован честной сетью и атакующий сформировал блок, победа злоумышленника, гонка завершена, вероятность наступления события $p \cdot q = 0,21$;

[1] В этом случае победа будет засчитана только после формирования $N = 1$ блока честной сетью

2) вторая попытка (шестнадцать возможных исходов, рассматриваем только те случаи, когда после первой попытки исход гонки не определен)

- (в первой попытке блок сформирован честной сетью и атакующий не сформировал блок):
  - во второй попытке блок сформирован честной сетью и атакующий не сформировал блок, злоумышленник проиграл в гонке, гонка завершена, вероятность наступления события $p \cdot (1-q) \cdot p \cdot (1-q) = 0{,}2401$;
  - во второй попытке блок не сформирован честной сетью и атакующий сформировал блок, победа злоумышленника, гонка завершена, вероятность наступления события $p \cdot (1-q) \cdot (1-p) \cdot q = 0{,}0441$;
  - во второй попытке блок не сформирован честной сетью и атакующий не сформировал блок, гонка продолжается, вероятность наступления события $p \cdot (1-q) \cdot (1-p) \cdot (1-q) = 0{,}1029$;
  - во второй попытке блок сформирован честной сетью и атакующий сформировал блок, злоумышленник проиграл в гонке, гонка завершена, вероятность наступления события $p \cdot (1-q) \cdot p \cdot q = 0{,}1029$;
- (в первой попытке блок не сформирован честной сетью и атакующий не сформировал блок):
  - во второй попытке блок сформирован честной сетью и атакующий не сформировал блок, гонка продолжается, вероятность наступления события $(1-p) \cdot (1-q) \cdot p \cdot (1-q) = 0{,}1029$;
  - во второй попытке блок не сформирован честной сетью и атакующий сформировал блок, победа злоумышленника[2], вероятность наступления события $(1-p) \cdot (1-q) \cdot (1-p) \cdot q = 0{,}0189$;
  - во второй попытке блок не сформирован честной сетью и атакующий не сформировал блок, гонка продолжается, вероятность наступления события $(1-p) \cdot (1-q) \cdot (1-p) \cdot (1-q) = 0{,}0441$;
  - во второй попытке блок сформирован честной сетью и атакующий сформировал блок, победа злоумышленника, гонка завершена, вероятность наступления события $(1-p) \cdot (1-q) \cdot p \cdot q = 0{,}0441$.

Таким образом, в модели независимых игроков злоумышленнику удастся победить за две попытки с вероятностью $0{,}09 + 0{,}21 + 0{,}0441 + 0{,}0189 + 0{,}0441 = 0{,}4071$, что отличается от вероятности, рассчитанной для модели разорения игрока.

С помощью моделирования проведем вычислительный эксперимент и эмпирически оценим вероятности выигрыша злоумышленника у честной сети при различных моделях формирования цепочки блоков.

4. Моделирование вычислительного эксперимента

На первом этапе протестируем вероятности формирования N блоков ровно за t попыток и вероятность формирования N блоков при проведении t испытаний.

4.1. Вероятность формирования цепочки блоков заданной длины

На первом этапе протестируем вероятность формирования блока ровно за t попыток.

В качестве входных параметров будем использовать:

q – вероятность успешно сформировать блок злоумышленником при каждом испытании. Вероятность зависит от имеющихся у злоумышленника вычислительных возможностей (т.е. пропорциональна хешрейту злоумышленника);

p – вероятность успешно сформировать блок честными участниками при каждом испытании (пропорционально хешрейту честной сети). При моделировании будем полагать, что

[2] В этом случае победа будет засчитана только после формирования $N = 1$ блока честной сетью

$p = 1 - q$, так как такая взаимосвязь лежит в основе модели разорения игрока. В общем случае для модели независимых игроков условие $p = 1 - q$ может не выполняться;

N – количество блоков в сети, после которых сделка считается подтвержденной;

t – номер текущей попытки.

В программной среде создадим процесс, итеративно пытающийся сформировать блоки. Каждое испытание происходит по следующему правилу:

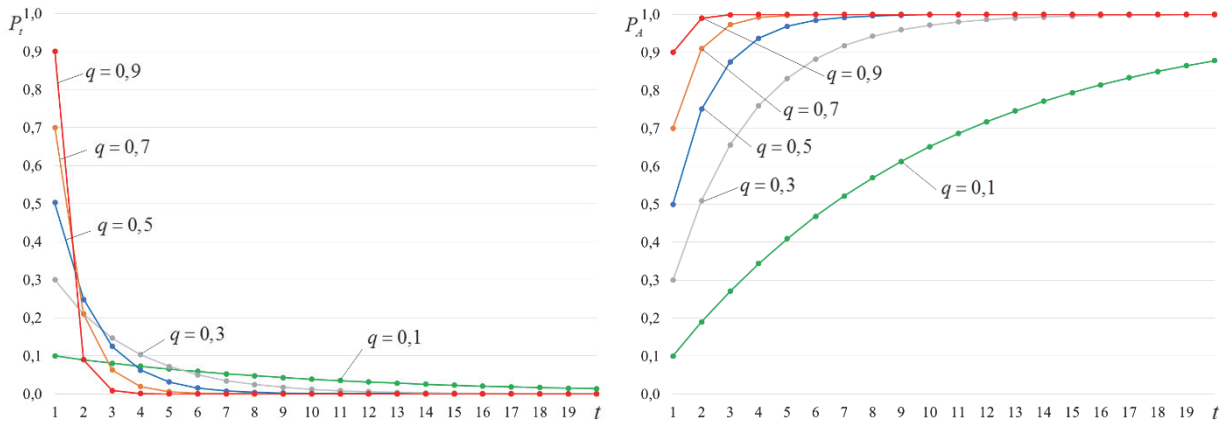
– генерируем псевдослучайное число (используем реализацию на основе Вихря Мерсенна) в интервале $[0,1]$ (в программной реализации используется минимальный шаг генерации $5.4 \cdot 10^{-20}$, что позволяет тестировать $q \geq 10^{-19}$);

– сравниваем сгенерированное случайное число со значением q;

– если сгенерированное число $\leq q$, то считаем, что генерация блока прошла успешной и увеличиваем счетчик сформированных потоком блоков ($k\_блок1$) на единицу. Проверяем $k\_блок1 = N$, если сформировали необходимо количество блоков, то увеличиваем массив $Mass1[t]$ на единицу, что соответствует удачной попытке сформировать цепочку блоков нужной длины на t-й попытке;

Для достижения заданной точности испытания проводим N_{test} раз (выбор N_{test} описан ниже). По окончании всех испытаний нормируется результат (массив $Mass1[t]$) по общему количеству испытаний и таким образом получаем эмпирическое распределение вероятности формирования N блоков от количества проведенных попыток P_t. Суммируя все полученные вероятности от 1 до заданного t получаем эмпирическую функцию распределения вероятности формирования блока P_A.

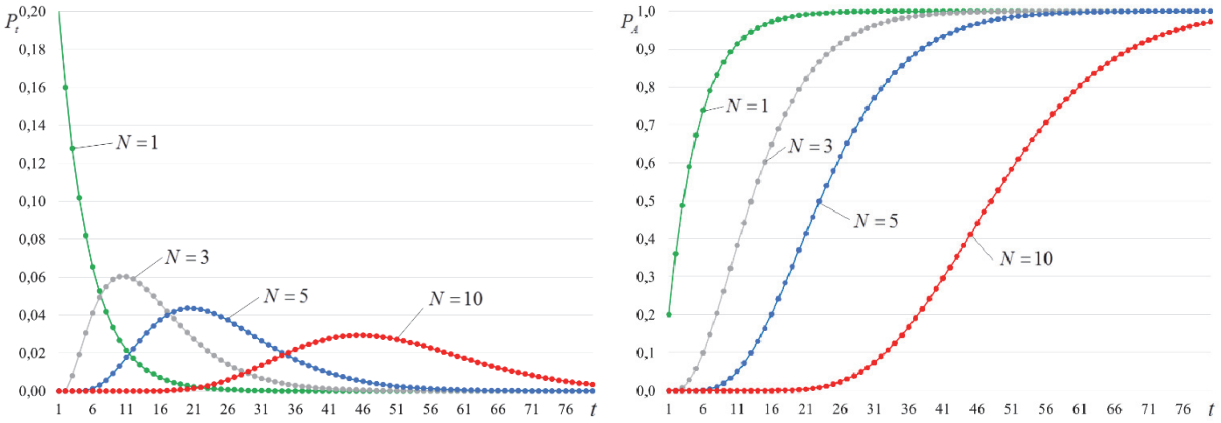
Результаты испытаний приведены на рис. 1 (точки). Сплошные линии соответствуют отрицательному биномиальному распределению и его функции распределения для тех же вероятностей.



а *б*

Рис. 1. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования блока при каждом испытании (линиями отображены расчетные значения, соответствующие отрицательному биномиальному распределению, точки – экспериментальные данные)

Построим вероятность формирования цепочки из N блоков. На рис. 2 представлены полученные аналогичные вероятности, но для фиксированного значения q и разного числа N.



а *б*

Рис. 2. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования цепочки из N блоков при $q = 0,2$ (линиями отображены расчетные значения, точки соответствуют экспериментальным данным)

Как видим, значения, полученные вычислительным моделированием, хорошо согласуются с отрицательным биномиальным распределением.

На втором этапе будем моделировать двух конкурирующих участников.

Будем исследовать две модели соревнования (гонки) злоумышленника с честной сетью по формированию цепочки блоков:

– модель разорения игрока;

– модель независимых игроков.

4.2. Модель разорения игрока

В программной среде создадим процесс (соответствующий злоумышленнику), итеративно пытающийся сформировать блоки. Каждое испытание происходит по следующему правилу:

– генерируем случайное число в интервале $[0,1]$;

– сравниваем полученное число с q;

– если сгенерированное число $\le q$, то считаем, что генерация блока прошла успешной и увеличиваем счетчик сформированных злоумышленником блоков ($k\_блок1$) на единицу;

– если сгенерированное число $> q$, то считаем, что блок сгенерирован честной сетью и увеличиваем счетчик сформированных ею блоков ($k\_блок2$) на единицу. Проверяем $k\_блок2 \ge N$, если да, то проверяем: сформировал ли злоумышленник необходимой длины цепочку:

- если злоумышленнику также удалось сформировать необходимое количество блоков (т.е. $k\_блок1 \ge k\_блок2$) то увеличиваем $Mass1[t]$ на единицу, что соответствует удачной попытке злоумышленника сформировать цепочку блоков нужной длины за t попыток (злоумышленник выиграл гонку). Завершаем испытание;
- если злоумышленник еще не сформировал необходимое количество блоков (т.е. $k\_блок1 < k\_блок2$), то продолжаем испытание;

– если $k\_блок2 = N + n_{\max}$, заканчиваем испытание, присваивая победу честной сети (увеличиваем массив $Mass2[t]$ на единицу).

В вычисленных экспериментах положим $n_{\max} = 1000$ (что соответствует практически неограниченным ресурсам злоумышленника). При выборе $n_{\max}$ мы упомянем работу [11],

в которой утверждается, что для $q < 0.45$ выбор значение $n_{\max} = 35$ практически не влияет на результат, кроме того, данный вопрос будет рассмотрен ниже.

4.3. Модель независимых игроков

В программной среде создадим два независимых процесса (первый процесс, соответствующий злоумышленнику, второй – честным пользователям), итеративно пытающихся сформировать блоки. Каждое испытание происходит по следующему правилу:

– генерируем случайное число в интервале $[0,1]$;

– сравниваем полученное число с q;

– если сгенерированное число $\le q$, то считаем, что генерация блока прошла успешной и увеличиваем счетчик сформированных первым потоком блоков ($k\_блок1$) на единицу;

– генерируем случайное число в интервале $[0,1]$;

– сравниваем полученное число с p;

– если сгенерированное число $\le p$, то считаем, что генерация блока прошла успешной и увеличиваем счетчик сформированных вторым потоком блоков ($k\_блок2$) на единицу. Проверяем $k\_блок2 \ge N$, если да, то проверяем: сформировал ли первый поток необходимой длины цепочку:

- если и злоумышленнику также удалось сформировать необходимое количество блоков (т.е. $k\_блок1 \ge k\_блок2$) то увеличиваем $Mass1[t]$ на единицу, что соответствует удачной попытке сформировать цепочку блоков нужной длины злоумышленником за t попыток (злоумышленник выиграл гонку). Завершаем испытание;
- если злоумышленник еще не сформировал необходимое количество блоков, то продолжаем испытание;

– если $k\_блок2 = N + n_{\max}$ заканчиваем испытание присваивая победу честной сети (увеличиваем массив $Mass2[t]$ на единицу).

4.4. Обеспечение точности и достоверности результатов моделирования

С помощью имитационного моделирования точное значение случайной величины (обозначим ее Θ) определить нельзя, так как число реализаций модели ограничено. При конечном числе реализаций модели определяется приблизительное значение заданной характеристики. Обозначим это приближение Θ^*. Приблизительное значение называют оценкой соответствующей характеристики [15, 16].

Точностью оценки характеристики Θ^* называют величину ε относительно

$$\left|\Theta^* - M\left[\Theta\right]\right| < \varepsilon,$$

где $M\left[\Theta\right]$ – математическое ожидание случайной величины [15, 16].

Величина ε представляет собой абсолютное значение ошибки в определении значения искомой характеристики.

Достоверностью оценки характеристики Θ^* называют вероятность α того, что заданная точность достигается [15, 16]:

$$P\left(\left|\Theta^* - M\left[\Theta\right]\right| < \varepsilon\right) = \alpha.$$

Достоверность характеризует повторяемость, устойчивость эксперимента и трактуется так: если для оценки $M\left[\Theta\right]$ использовать величину Θ^*, то в среднем на каждые 1000 использований данного правила в $1000 \cdot \alpha$ случаев величина Θ^* будет отличаться от $M\left[\Theta\right]$ на величину меньше ε.

В ряде случаев целесообразно использовать относительную точность

$$d = \varepsilon / M\left[\Theta\right].$$

В этом случае достоверность оценки имеет вид

$$P\left(\left|\frac{\Theta^* - M\left[\Theta\right]}{M\left[\Theta\right]}\right| < d\right) = \alpha .$$

Если принять предположение относительно нормального распределения случайной величины[3], тогда функциональная связь между относительной точностью и достоверностью с количеством реализаций N_{test} имеет вид [15]:

$$N_{test} = \frac{t_\alpha{}^2 (1-P)}{Pd^2},$$

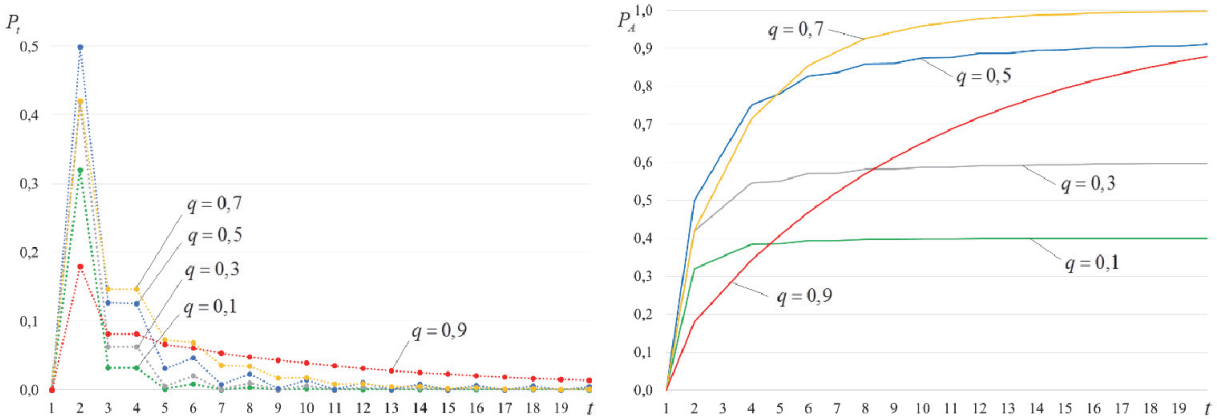
где t_α – аргумент функции Лапласа $t_\alpha = \Phi_0^{-1}\left(\frac{\alpha}{2}\right)$, интеграл Лапласа табулированный, следовательно, задаваясь значением достоверности α, можем определить t_α.

Из последней формулы следует, что при определении оценок малых вероятностей с приемлемой точностью необходимо выполнить очень большое число реализаций модели. При отсутствии высокопроизводительного компьютера применения статистического моделирования становится проблематичным.

Для проведения экспериментальных исследований были выбраны $\alpha = 0.99$ и $d = 0.01$, значение N_{test} рассчитывалось по приведенной выше формуле.

5. Результаты вычислений

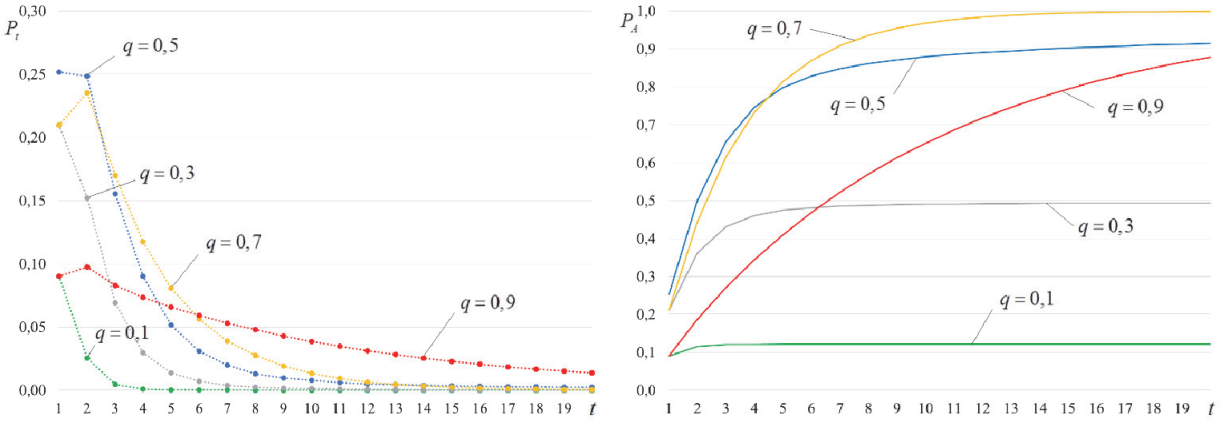
С использованием рассмотренных моделей были получены эмпирические оценки вероятности удачного формирования злоумышленником цепочки блоков при разных значениях q и N. На рис. 3 – 8 приведены полученные результаты в зависимости от номера попытки для каждого испытания, а также представлены соответствующие функции распределения вероятности в зависимости от количества попыток для каждого испытания.



а *б*

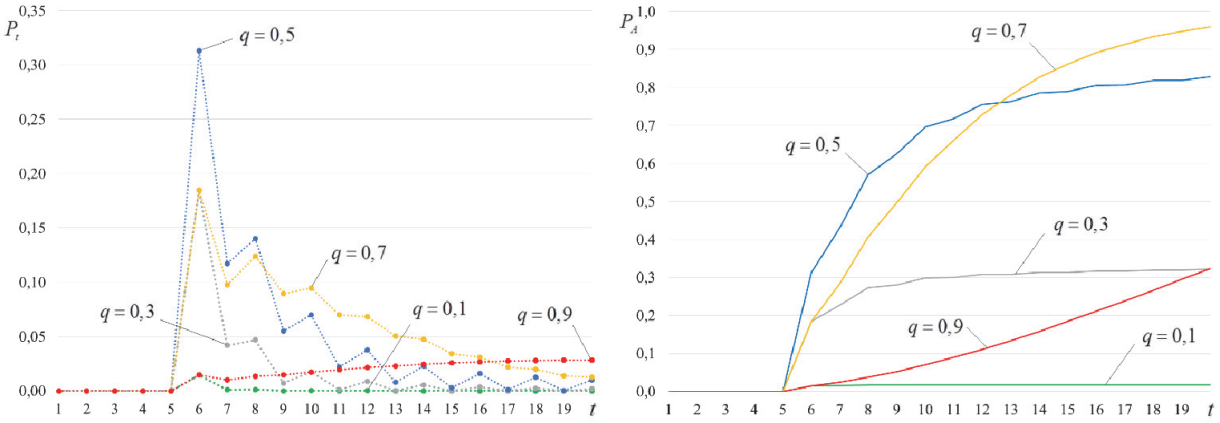
Рис. 3. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования цепочки для $N = 1$ подтверждений злоумышленником при участии двух конкурирующих субъектах (модель разорения игрока)

[3] В силу центральной предельной теоремы для большого числа испытаний биномиальное распределение хорошо аппроксимируется нормальным распределением [15, 16]



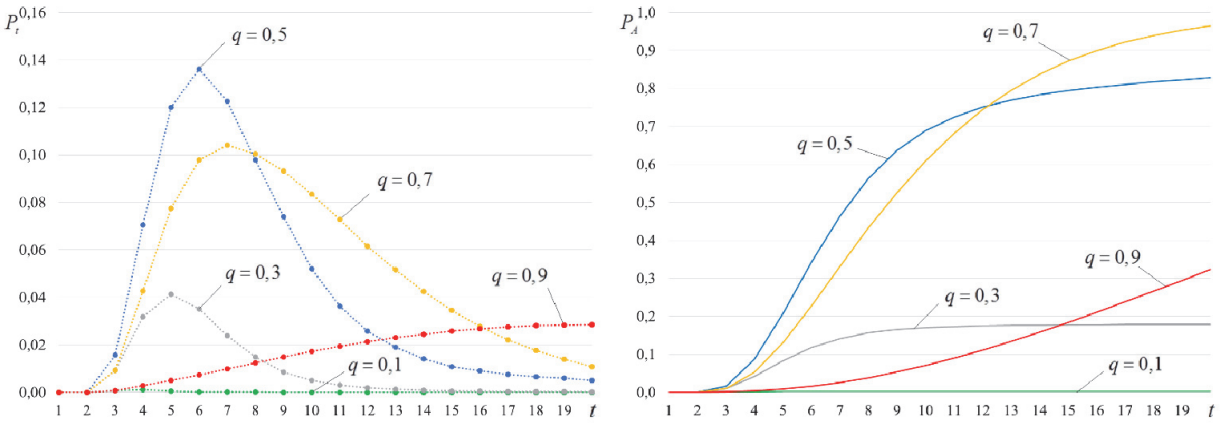
а *б*

Рис. 4. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования цепочки для $N = 1$ подтверждений злоумышленником при участии двух конкурирующих субъектах (модель независимых игроков)



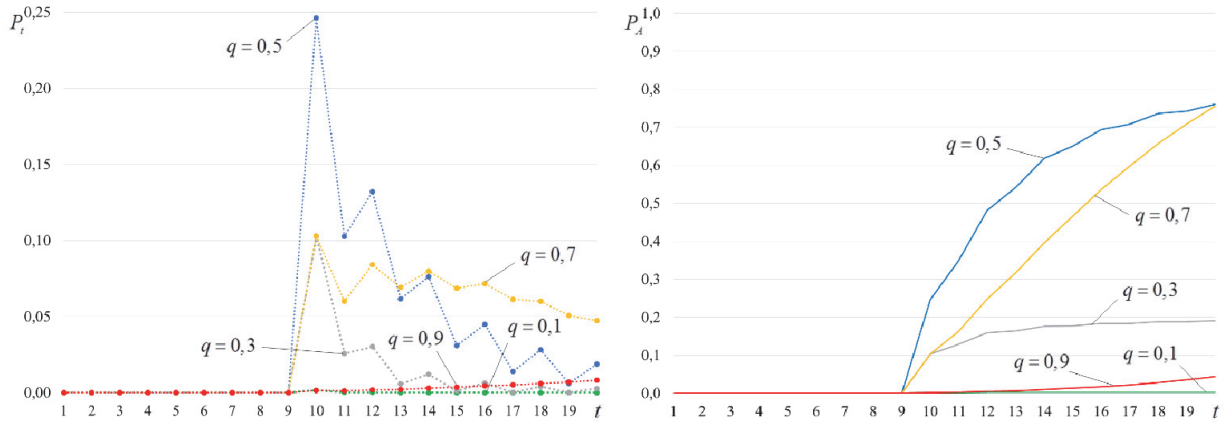
а *б*

Рис. 5. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования цепочки для $N = 3$ подтверждений злоумышленником при участии двух конкурирующих субъектах (модель разорения игрока)



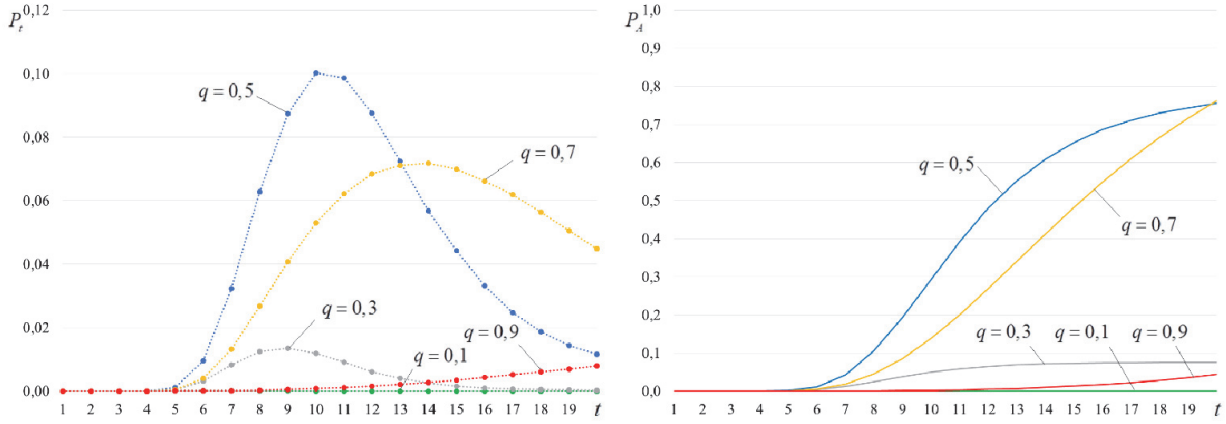
а *б*

Рис. 6. Функция вероятности (*а*) и функция распределения вероятности (*б)* формирования цепочки для $N = 3$ подтверждений злоумышленником при участии двух конкурирующих субъектах (модель независимых игроков)



а *б*

Рис. 7. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования цепочки для $N = 5$ подтверждений злоумышленником при участии двух конкурирующих субъектах (модель разорения игрока)

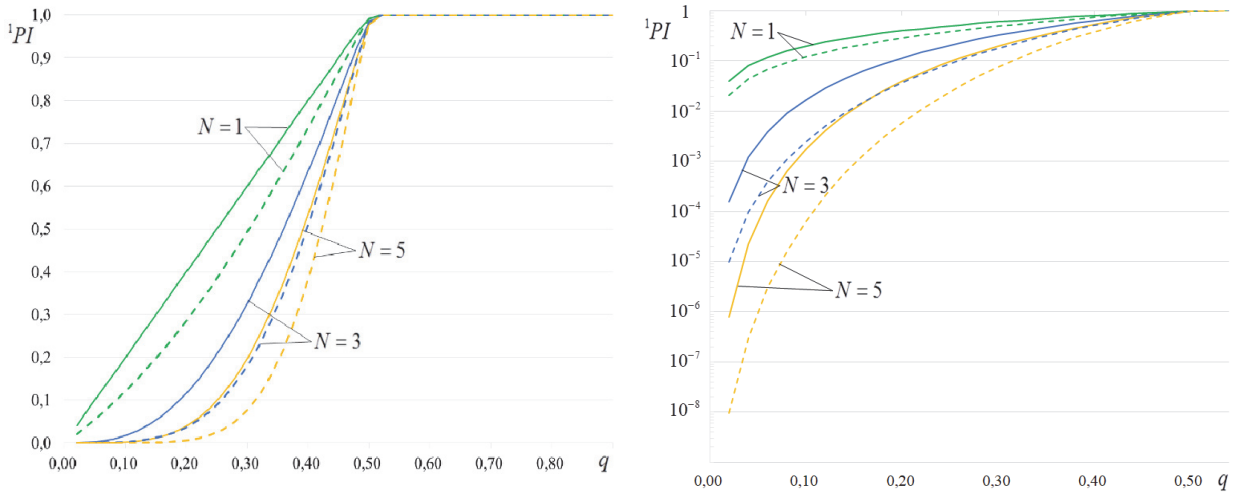


а *б*

Рис. 8. Функция вероятности (*а*) и функция распределения вероятности (*б*) формирования цепочки для $N = 5$ подтверждений злоумышленником при участии двух конкурирующих субъектах (модель независимых игроков)

Просуммировав описанные вероятности по всем возможным испытаниям, то есть для всех $t = 1, 2, 3, \ldots$, получим интегральную (или общую) вероятность успешного формирования альтернативной цепочки блоков для N подтверждений злоумышленником (^1PI).

Для приведенных примеров интегральная вероятность успешного формирования злоумышленником цепочки для N подтверждений приведена на рис. 9. Для удобства анализа полученных данных один и тот же результат приведен в обычной шкале (хорошо иллюстрирует поведение кривых при $q > 0,2$) и в логарифмическом масштабе (для иллюстрации кривых при $q < 0,2$). В этих и последующих графиках изменение значения q проводилось с шагом 0,02.



а *б*

Рис. 9. Интегральная вероятность успешного формирования злоумышленником цепочки блоков для N подтверждений при участии двух конкурирующих субъектах (экспериментальные данные). Сплошная линия – модель разорения игрока, пунктиром – модель независимых игроков. *а* – обычная шкала, *б* – логарифмическая шкала

Как видим, результаты различных моделей значительно отличаются друг от друга. Рассмотрим относительную ошибку моделирования (для двух рассмотренных моделей), определяемую как

$$\frac{{}^1PI_{мри} - {}^1PI_{мни}}{{}^1PI_{мри}} \cdot 100\%,$$

где ${}^1PI_{мри}$ – интегральная вероятность, рассчитанная на основе модели разорения игрока; ${}^1PI_{мни}$ – интегральная вероятность, рассчитанная на основе модели независимых игроков,

При веденном обозначении значения относительной ошибки моделирования приведены в табл. 1.

Таблица 1

Значение относительной ошибки в результате применения различных моделей (на основе разорения игрока и независимых игроков)

	$q = 0,01$	$q = 0,2$	$q = 0,4$
$N = 1$	48 %	29 %	8 %
$N = 3$	94 %	68 %	20 %
$N = 5$	99 %	85 %	28 %

Как видим из таблицы, две рассмотренные модели атаки двойной траты (модель разорения игроков и модель независимых игроков) дают различные оценки вероятности выигрыша гонки злоумышленником (успеха атаки). По мере увеличения длины цепочки блоков N расхождение увеличивается (относительная ошибка моделирования достигает 100 %). Это наблюдается для различных вероятностей q (т.е. для различных соотношений хешрейтов злоумышленника и честной сети).

Следует отметить, что результаты, полученные на основе модели разорения игрока, соответствуют (в пределах заданной достоверности и выбранной относительной точности) аналитическим результатам, полученным на основе формул М. Розенфельда (см. выражение 1 и рис. 4 из [5]. Отличие наблюдается только в точке $q = 0,5$, где относительная ошибка между

экспериментальными и аналитическими результатами составила 1,7 % для $N = 3$ и 2,2 % для $N = 5$, что связанно с ограничением в $n_{\max} = 1000$ блоков.

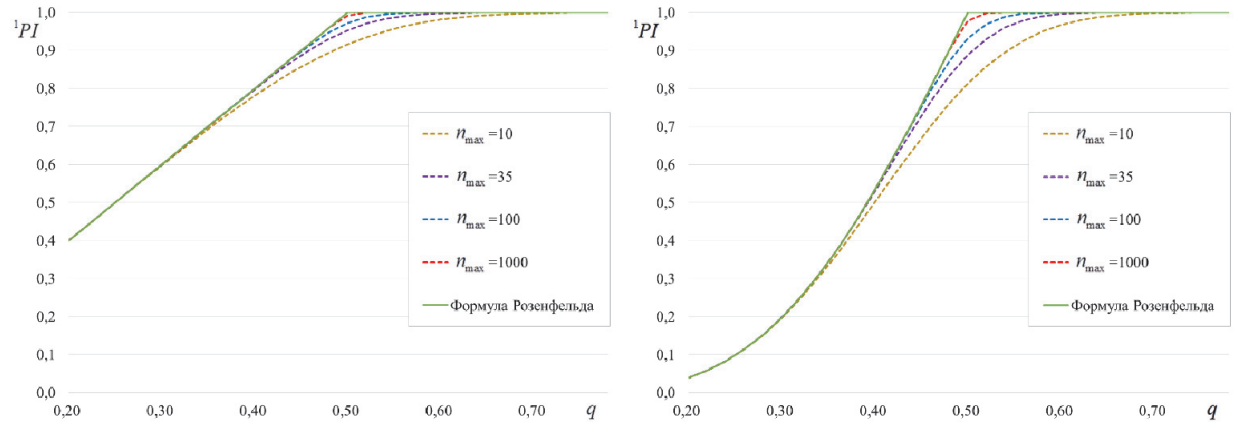
Как было показано в работе [11], результат имеет отличия при разных значениях $n_{\max}$. Проанализируем данный вопрос подробнее.

6. Влияние $n_{\max}$ на вероятность победы злоумышленника

Учитывая, что поддержка гонки злоумышленником постоянно требует определенных финансовых затрат от злоумышленника, то гонка только теоретически может продолжаться бесконечно. В реальных обстоятельствах злоумышленнику будет невыгодно продолжать гонку и затрачивать на ее поддержание больше ресурсов, чем он сможет себе вернуть, удачно проведя атаку двойной траты, или имеет в своем распоряжении. Другой вариант, если злоумышленник в состоянии формировать определенное количество блоков на протяжении большого промежутка времени, то ему может быть экономически выгодней их публиковать по правилам сети, получая за это награду, чем пытаться извлечь выгоду из нечестного (не соответствующего правилам сети) поведения. Если злоумышленник отстал в гонке с честной сетью на значительное количество блоков, то, как показано выше, его шансы победить значительно снижаются, и ему уже нет смысла продолжать попытки до бесконечности.

При всех рассмотренных вариантах значение $n_{\max}$ есть число конечное. Рассмотрим его влияние на вероятность победы злоумышленника.

В качестве иллюстрации на рис. 10 приведены графики экспериментальных значений, полученных в соответствии с моделью разорения игрока, для $n_{\max} = 10, 35, 100, 1000$ и разных $N = 1$ и 5, а также проводится сравнение с теоретическими результатами, полученными М. Розенфельдом.



а *б*

Рис. 10. Интегральная вероятность успешного формирования злоумышленником цепочки блоков для $N = 1$ (*а*) и $N = 5$ (*б*) подтверждений при участии двух конкурирующих субъектов (пунктир – экспериментальные данные). Модель разорения игрока

Как видно из приведенных результатов, увеличение $n_{\max}$ приближает полученные эмпирические данные к аналитическим результатам М. Розенфельда [5]. С уменьшением вероятности теоретические результаты, полученные М. Розенфельдом, хорошо аппроксимируются при небольших $n_{\max}$.

Относительная ошибка между теоретическими и экспериментальными результатами близка к значению $q = 0,5$ и составляет более 0,1 % в следующих диапазонах:

для $N = 1$:

– от $0,28 \le q \le 0,72$ при $n_{\max} = 10$;

– от $0,4 \le q \le 0,6$ при $n_{\max} = 35$;

– от $0,44 \le q \le 0,56$ при $n_{\max} = 100$;

– $q = 0,5$ при $n_{\max} = 1000$;

для $N = 5$:

– от $0,26 \le q \le 0,70$ при $n_{\max} = 10$;

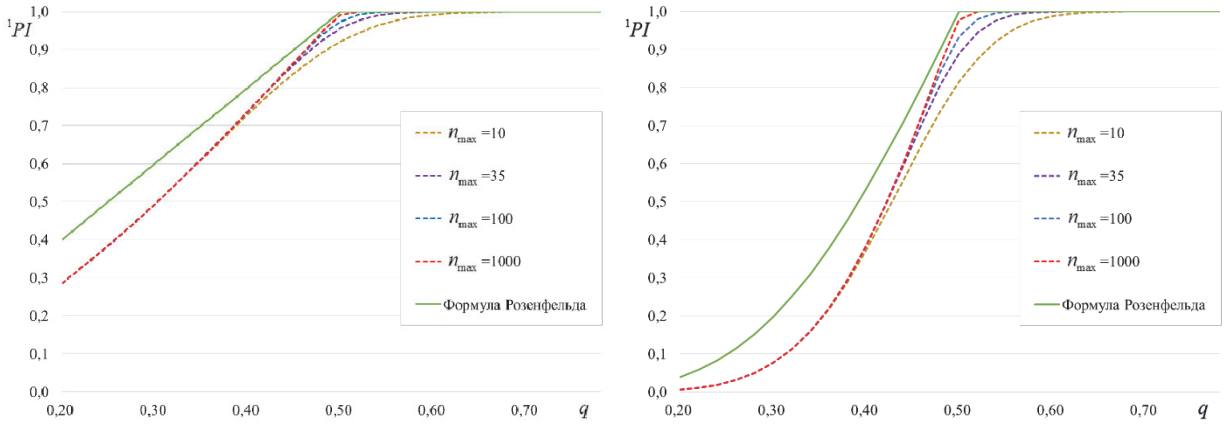
– от $0,38 \le q \le 0,62$ при $n_{\max} = 35$;

– от $0,42 \le q \le 0,54$ при $n_{\max} = 100$;

– от $0,48 \le q \le 0,5$ при $n_{\max} = 1000$;

Сходимость результатов экспериментов с теоретическими расчетами по известным аналитическим выражениям подтверждает адекватность и обоснованность результатов исследований.

На рис. 11 приведены экспериментальные результаты, полученные в соответствии с моделью независимых игроков при тех же параметрах ($n_{\max} = 10, 35, 100, 1000$; $N = 1, 5$). Для наглядности оставлен теоретический результат, полученный М. Розенфельдом.



а *б*

Рис. 11. Интегральную вероятность успешного формирования злоумышленником цепочки блоков для $N = 1$ (*а*) и $N = 5$ (*б*) подтверждений при участии двух конкурирующих субъектах (пунктир – экспериментальные данные). Модель независимых игроков.

Как видим, характер влияния $n_{\max}$ на полученный результат сохраняется и для модели независимых игроков. Однако сопоставление полученных результатов для различных моделей подтверждает тезис о расхождении оценок вероятностей успешной атаки двойной траты.

7. Выводы

Проведен критический анализ известных работ по оценке вероятностей двойной траты в протоколе консенсуса «Proof of work». Показано наличие неточностей и необоснованных допущений в известных работах С. Накамото [4] и М. Розенфельда [5]. В частности, показано, что базовые предположения о вероятностном пространстве (множество элементарных исходов и вероятности их наступления) в используемой модели разорения игроков (с двумя элементарными исходами) не соответствуют реальным процессам, протекающим при установлении консенсуса «Proof of work».

Для теоретической оценки вероятности успешной атаки двойной траты предложено использовать модель независимых игроков с четырьмя элементарными исходами. Эта модель описывает реальный вероятностный процесс в блокчейн-системе при установлении консенсуса на основе алгоритма «Proof of work», когда каждый участник (злоумышленник и честная сеть) независимо формируют блоки с вероятностями, пропорциональными своему хешрейту (своим вычислительным возможностям).

Проведено сравнение результатов, полученных с помощью вычислительного моделирования атаки двойной траты на основе модели разорения игрока и модели независимых игроков. Сравнение проведено для разных возможностей злоумышленника (вероятности сформировать блок), различного количества сформированных блоков, после которых сделка считается подтвержденной, различной продолжительности гонки (количества блоков, на протяжении которых злоумышленник продолжает попытки догнать честную сеть). Показано значительное отличие (относительная ошибка модели до 99 %) результатов, полученных в вычислительном моделировании при использовании модели независимых игроков от модели разорения игрока.

Все эмпирические оценки получены для высокой точности (относительная ошибка не более 1 %) и достоверности (доверительная вероятность не менее 99 %).

Для подтверждения адекватности полученных результатов приведено сравнение эмпирических результатов с теоретическими расчетами по известным аналитическим соотношениям. Показано, что результаты вычислительного эксперимента для модели разорения игрока полностью совпадают (в пределах заданной достоверности и относительной точности) с аналитическим результатом, приведенным в работе М. Розенфельда [5].

На основе полученных результатов можно утверждать об ошибочности использования модели разорения игрока для оценки вероятности успешной атаки двойной траты на протокол консенсуса «Proof of work».

Список литературы:

1. The Double Spending Problem and Cryptocurrencies. Banking & Insurance Journal. Social Science Research Network (SSRN). Accessed 24 December 2017. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3090174
2. Mark Ryan. Digital Cash // School of Computer Science, University of Birmingham. Retrieved 2017-05-27. https://www.cs.bham.ac.uk/~mdr/teaching/modules06/netsec/lectures/DigitalCash.html
3. Varshney, Neer (2018-05-24). Why Proof-of-work isn't suitable for small cryptocurrencies // Hard Fork. Retrieved 2018-05-25. https://thenextweb.com/hardfork/2018/05/24/proof-work-51-percent-attacks/
4. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System / Satoshi Nakamoto, 2009. 9 с.
5. Rosenfeld M. Analysis of hashrate-based double-spending / Meni Rosenfeld, 2014. 13 с.
6. Carlos Pinzón, Camilo Rocha. Double-spend Attack Models with Time Advantange for Bitcoin // Electronic Notes in Theoretical Computer Science. Volume 329, 9 December 2016, Pages 79-103 https://doi.org/10.1016/j.entcs.2016.12.006
7. Kaidalov D.S., Kovalchuk L.V., Nastenko A.O., Rodinko M.Yu., Shevtsov O.V., Oliynykov R.V. Comparison of block expectation time for various consensus algorithms // Radio Electronics, Computer Science, Control. 2018. № 4. PP. 159- 171 DOI 10.15588/1607-3274-2018-4-15
8. Azzolini D., Riguzzi F., Lamma E., Bellodi E., Zese R. Modeling Bitcoin Protocols with Probabilistic Logic Programming http://ceur-ws.org/Vol-2219/paper6.pdf
9. Kevin Liao, Jonathan Katz. Incentivizing Double-Spend Collusion in Bitcoin. 2017. https://www.cs.umd.edu/~gasarch/reupapers/katzbitcoin16.pdf
10. Ковальчук Л.В. Основні визначення у галузі блокчейну та детальний аналіз результатів Накамото-Розенфельда-Грунспана про імовірність атаки подвійної витрати. Звіт про НДР (проміжний). Харків : АТ ІІТ. 36 с.
11. Pinar Ozisik., Brian Neil Levine. An Explanation of Nakamoto's Analysis of Double-spend Attacks https://arxiv.org/pdf/1701.03977.pdf
12. Apostolaki M. Hijacking Bitcoin: Routing Attacks on Cryptocurrencies / M. Apostolaki, A. Zohar, L. Vanbever. San Jose, CA , USA, 2017. 18 с.
13. Grunspan C., Pérez-Marco R. Double spend races. 2017. hal-01456773 https://hal.archives-ouvertes.fr/hal-01456773
14. W. Feller. An Introduction to Probability Theory and its Applications: Volume I, volume 3. John Wiley & Sons London-New York-Sydney-Toronto, 1968
15. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. Москва : Наука, 1969. 512 с.
16. Вентцель Е.С. Теория вероятности. Москва : Наука, 1969. 576с.

Харьковский национальный
университет имени В.Н. Каразина;
АО «Институт информационных технологий», Харьков

Поступила в редколлегию 27.08.2019