

## АЛГОРИТМ ОПРЕДЕЛЕНИЯ ИСТОЧНИКА ФРАГМЕНТИРОВАННЫХ СООБЩЕНИЙ

М. О. ТАНЫГИН

*Юго-Западный государственный университет, 305040, Курск, Россия  
E-mail: tanygin@yandex.ru*

Рассматривается проблема снижения информационной избыточности при обмене данными в сетях с большим радиусом действия и низким энергопотреблением за счет уменьшения размеров дополнительных служебных полей, по которым происходит определение источника информационного пакета и его порядкового номера в последовательности пакетов, составляющих единое сообщение. Для повышения энтропии этих дополнительных служебных полей предлагается записывать в них хеш-алгоритм, сформированный из данных других информационных пакетов источника и уникального идентификатора источника. Описан формализованный алгоритм анализа приемником поступающих информационных пакетов. Сформулирована проблема возникновения ошибки определения порядкового номера фрагмента в едином сообщении. На основе математического аппарата теории вероятностей получены численные характеристики такой ошибки, произведена оценка ее влияния на информационную избыточность протоколов, использующих рассматриваемый алгоритм. Приведены зависимости между длиной дополнительного поля хеша и количеством фрагментов, на которые разделено передаваемое сообщение, определен диапазон значений, при котором достигается минимальная информационная избыточность.

**Ключевые слова:** *определение источника сообщений, теория вероятностей, хеш, нарушение порядка следования пакетов, параметры алгоритма, информационная избыточность*

Развитие технологии так называемого „интернета вещей“ (IoT — Internet of Things) подтолкнуло создателей телекоммуникационного оборудования к разработке новых стандартов и протоколов взаимодействия устройств вычислительной техники [1—3]. Характерной особенностью таких устройств является низкое энергопотребление (до 10 лет от стандартного элемента питания) и низкая скорость передачи данных (до 300 бит/с) при большом радиусе действия [4, 5]. Следствие подобных ограничений — небольшой размер информационного пакета, передаваемого между абонентами на физическом уровне (до 18 байт с учетом дополнительных служебных полей) [4]. Учитывая необходимость поддерживать сложные топологии сетей, объединяющих взаимодействующие устройства, а также низкую защищенность протокола криптографическими средствами [5], необходимо осуществлять контроль источника каждого информационного пакета на физическом уровне. Спецификацией действующих и планируемых к внедрению протоколов предусматривается идентификация каждого пакета по 3-4-байтовому идентификатору отправителя. В сочетании с несколькими байтами поля CRC-кодов, служащих для обнаружения и устранения ошибок, объем дополнительной служебной информации, передаваемой в пакете, может варьироваться от 50 до 70 % [1—3]. В этой связи становится актуальной задача контроля источника пакета с помощью методов и алгоритмов, использующих для контроля меньшее количество битов информации в служебных полях. На практике это позволило бы добиться, прежде всего, снижения нагрузки на радиопередающие устройства, работающие на данных протоколах, и, как следствие, повышения срока службы их элементов питания.

В работах [6—8] рассматриваются протоколы, в которых размер блока данных не превышает нескольких десятков и даже нескольких байтов, что обуславливает актуальность вопросов информационной избыточности. В основу протоколов положен принцип принадлежности отдельного информационного пакета (ИП) некоей группе, ассоциируемой с конкретными отправителем. Несмотря на то, что описанные системы имеют ряд недостатков, как то: низкая энтропия в дополнительных служебных полях [9] и большой объем априорной информации, которую приемник должен хранить о каждом передающем устройстве, подход, основанный на групповой проверке пакетов, представляется наиболее целесообразным для контроля источника в условиях ограниченности размера служебных полей.

Для объединения информационных пакетов в структурированные последовательности или цепочки и последующего их ассоциирования с определенным источником в настоящей статье предлагается использовать метод, основанный на сравнении значений хеш-функций, полученных из данных предыдущих пакетов [10—12]. Таким образом, достигается наиболее полное использование каждого бита дополнительного сервисного поля с точки зрения содержащегося в нем количества информации как меры устранения неопределенности. Приемник анализирует значения хеш-функций полученных пакетов и определяет их место на основании взаимного совпадения. Функция  $F^R$  определения позиции элемента  $s_i$  в множестве выделенных ИП основывается на вычислении значений хеш-функций предыдущих ИП и сравнении их с информацией самого пакета. Значение хеша одного или нескольких (в зависимости от реализации) предыдущих пакетов вводится в структуру каждого из ИП  $s_1, \dots, s_n$ , образующих сообщение  $S$  источника. Взаимное расположение ИП определяется в соответствии с априорной информацией о конкретном источнике данных [13, 14]:

$$\forall s_i^{\text{hash}} \subset s_i, i = \overline{1, \dots, n}, s_i^{\text{hash}} = f(S, S^{\text{key}}), \quad (1)$$

где  $s_i^{\text{hash}}$  — часть, формируемая источником сообщения и содержащая информацию (хеш предыдущих пакетов) о месте пакета  $s_i$  в сообщении  $S$ .

Задача определения принадлежности конкретного ИП источнику и сообщению формулируется следующим образом. Пусть имеется множество  $U$  информационных пакетов, поступивших в приемник. На первом этапе к каждому из этих пакетов применяются операции выделения стартового  $f^{\text{start}}$  и конечного  $f^{\text{stop}}$  пакетов:

$$\begin{aligned} \exists! s_{\text{start}} \subset U, f^{\text{start}}(s_{\text{start}}, S^{\text{key}}) = 1, \\ \exists! s_{\text{stop}} \subset U, f^{\text{stop}}(s_{\text{stop}}, S^{\text{key}}) = 1. \end{aligned} \quad (2)$$

Затем к оставшимся ИП применяется операция  $f^{si}$  выделения информационной части и  $f^{sh}$  выделения части хеша:

$$\forall s \subset U, s^{\text{hash}} = f^{sh}(s, S^{\text{key}}), s^{\text{inf}} = f^{si}(s, S^{\text{key}}), \quad (3)$$

где  $s^{\text{hash}}$  — часть пакета  $s$ , содержащая значение хеш-функции, полученное из данных предыдущего пакета;  $s^{\text{inf}}$  — информационная часть пакета  $s$ .

На практике это может быть банальным кодированием с последующим разделением полученного слова на два подслова в соответствии с определенными  $S^{\text{key}}$  размерами обеих частей. Далее номер пакета в последовательности определяется следующими рекуррентными соотношениями:

$$\begin{aligned}
 R_0 &= F^R(s_{\text{start}}) = 0; \\
 R_i &= F^R(s_i) = R_{i-1} + 1 \Leftrightarrow s_i^{\text{hash}} = F_{\text{hash}}(s_{i-1}^{\text{inf}}); \\
 R_{n+1} &= n + 1 \Leftrightarrow s_{\text{stop}}^{\text{hash}} = F_{\text{hash}}(s_n^{\text{inf}}),
 \end{aligned}
 \tag{4}$$

где  $F_{\text{hash}}$  — хеш-функция, используемая источником для формирования, а приемником для проверки содержимого дополнительного сервисного поля (поля, содержащего значение);  $s_{\text{stop}}^{\text{hash}}$  — часть поля хеша стоп-пакета.

При этом мощность множества  $\tilde{u} = \{s_{\text{start}}, s_1, s_2, \dots, s_{n-1}, s_n, s_{\text{stop}}\}$  строго определена параметрами передачи. Если при расчете позиции пакетов формируется множество  $\tilde{u}'$ , начинающееся и заканчивающееся пакетами  $s_{\text{start}}$  и  $s_{\text{stop}}$ , но обладающее другой мощностью, то функция принадлежности для него будет иметь значение „ложь“. Сама функция принадлежности примет следующий вид:

$$\begin{aligned}
 B(\tilde{u}, S^{\text{key}}) &= 1 \Leftrightarrow \tilde{u}^{(0)} = s_{\text{start}} \wedge \tilde{u}^{(n+1)} = s_{\text{stop}} \wedge |\tilde{u}| = n + 2 \wedge \\
 \wedge f^{sh}(\tilde{u}^{(i)}, S^{\text{key}}) &= F_{\text{hash}}(f^{si}(\tilde{u}^{(i-1)}, S^{\text{key}})), \quad i = \overline{1, \dots, n+1} \wedge \\
 \wedge \tilde{u}^{(i)} \cap \tilde{u} - \tilde{u}^{(i)} &= \emptyset, \quad i = \overline{1, \dots, n},
 \end{aligned}
 \tag{5}$$

где  $\tilde{u}^{(0)}$  — стартовый пакет структурированного множества  $\tilde{u}$ ,  $\tilde{u}^{(i)}$  —  $i$ -й элемент множества  $\tilde{u}$ ,  $|\tilde{u}|$  — мощность множества  $\tilde{u}$ .

Функция принадлежности будет принимать значение „истина“ только тогда, когда мощность множества  $\tilde{u}$ , сформированного на основе рекуррентного правила (4), равна  $n+2$ , а структурированное множество номеров пакетов, определенных по тому же правилу, равно множеству целых чисел от 0 до  $n+1$ . Последний элемент конъюнкции в формуле (5) означает, что в структурированном множестве невозможно повторение ИП.

Приведенное описание функции принадлежности для метода на основе сравнения значений хеш-функций подразумевает истинность ее значения для одного структурированного подмножества  $\tilde{u}$  из всех возможных подмножеств множества  $U$ . В то же время возможны ситуации, когда из множества  $U$  можно сформировать два или более изоморфных ему структурированных множеств  $u_1, \dots, u_r$ , для которых функция принадлежности будет истинна. При этом сами множества не будут равны, так как в рассматриваемой задаче определения источника сообщений решается еще одна подзадача — восстановление исходного порядка ИП, образующих сообщение. Пример формирования двух изоморфных множеств с истинной функцией принадлежности приведен на рис. 1.

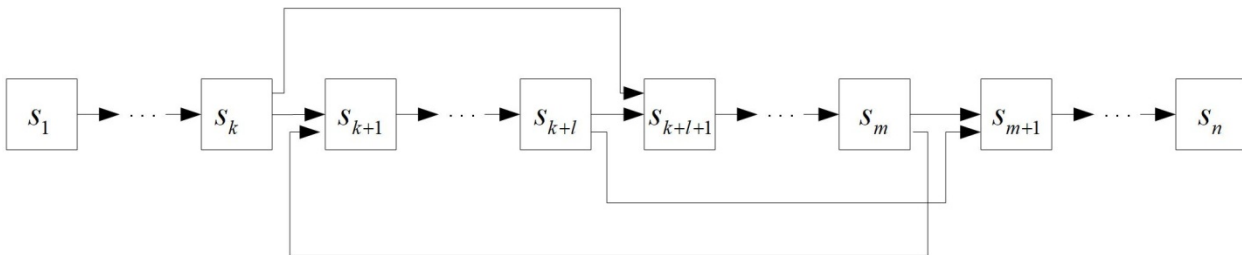


Рис. 1

Видно, что вследствие одновременного выполнения равенств

$$\left. \begin{aligned} s_{k+1}^{\text{hash}} &= F_{\text{hash}}(s_k^{\text{inf}}) = F_{\text{hash}}(s_m^{\text{inf}}) = s_{m+1}^{\text{hash}}; \\ s_{m+1}^{\text{hash}} &= F_{\text{hash}}(s_m^{\text{inf}}) = F_{\text{hash}}(s_{k+l}^{\text{inf}}) = s_{k+l+1}^{\text{hash}}; \\ s_{k+l+1}^{\text{hash}} &= F_{\text{hash}}(s_{k+l}^{\text{inf}}) = F_{\text{hash}}(s_k^{\text{inf}}) = s_{k+1}^{\text{hash}}, \end{aligned} \right\} s_{k+1}^{\text{hash}} = s_{m+1}^{\text{hash}} = s_{k+l+1}^{\text{hash}} \quad (6)$$

может быть сформировано два изоморфных множества, отличающиеся лишь порядком следования пакетов:

$$\begin{aligned} \tilde{y}_1 &= \{s_1, \dots, s_k, s_{k+1}, \dots, s_{k+l+1}, s_{k+l+1}, \dots, s_m, s_{m+1}, \dots, s_n\}, \\ \tilde{y}_2 &= \{s_1, \dots, s_k, s_{k+l+1}, \dots, s_m, s_{k+1}, \dots, s_{k+l}, s_{m+1}, \dots, s_n\}. \end{aligned} \quad (7)$$

Для определения вероятности возникновения подобной ошибки определим диапазоны изменения чисел  $k, l, m$ . Очевидно, что  $l$  может принимать значения в диапазоне от 1 до  $n-3$ ,  $k$  может варьироваться от 1 до  $n-2-l$ ,  $m$  — от  $k+l+1$  до  $n-1$ , где  $n$  — мощность множества  $\tilde{y}$  или число полученных приемником ИП.

Для нахождения итоговой вероятности рассмотрим ситуацию, когда  $k=1$ . Для того чтобы произошла ошибка, значение хеш-функции во втором пакете должно совпасть со значениями хеш-функций в минимум двух пакетах с номерами от 3 до  $n$ . Значения хеш-функций формируются независимо друг от друга, поэтому количество совпадений будет распределено по биномиальному закону. Соответственно вероятность ошибки для этого случая составит [15]

$$P_{\text{err}}(k=1) = \sum_{i=2}^{n-3} C_{n-3}^i (2^{-H})^i (1-2^{-H})^{n-3-i}, \quad (8)$$

где  $H$  — размер поля хеша ИП,  $C_{n-3}^i$  — число сочетаний из  $n-3$  по  $i$ .

В случае если среди  $n-3$  значений хеш-функций ни одно не совпадет со значением хеш-функции во втором пакете, следует перейти к рассмотрению ситуации для  $k=2$  с учетом того, что для  $n-4$  пакетов набор возможных значений хешей будет равен уже не  $2^H$ , а  $2^H-1$ , так как из общего количества будет исключено значение хеш-функции второго пакета. Тогда по аналогии с (8) ошибка для  $k=2$  определится как

$$P_{\text{err}}(k=2, \text{lh} = 1) = \sum_{i=3}^{n-4} C_{n-4}^i \left( \frac{1}{2^{-H}-1} \right)^i \left( 1 - \frac{1}{2^{-H}-1} \right)^{n-4-i}, \quad (9)$$

где  $\text{lh}$  — количество исключенных из рассмотрения вариантов содержимого поля  $s^{\text{hash}}$  ИП.

В случае если среди  $n-3$  результатов хеш-функций только один совпадет со значением поля  $s^{\text{hash}}$  второго пакета, следует перейти к рассмотрению ситуации для  $k=3$ , так как анализировать будем  $n-3$  пакетов (один пакет исключаем из рассмотрения, так как содержимое его поля  $s^{\text{hash}}$  совпало с содержимым данного поля второго пакета). При этом, как и в случае несовпадения содержимого полей значений хеш-функций, исключаем один вариант:

$$P_{\text{err}}(k=3, \text{lh} = 1) = \sum_{i=3}^{n-5} C_{n-5}^i \left( \frac{1}{2^{-H}-1} \right)^i \left( 1 - \frac{1}{2^{-H}-1} \right)^{n-4-i}. \quad (10)$$

Формулу для результирующей вероятности возникновения ошибки с учетом выражений (8)—(10), а также вероятностей совпадения значений хеш-функций соответствующее количество раз можно записать в следующем виде:

$$\begin{aligned}
P_{\text{err}}(k=1) = & \sum_{i=2}^{n-3} C_{n-3}^i \left(\frac{1}{2^{-H}}\right)^i \left(1 - \frac{1}{2^{-H}}\right)^{n-3-i} + \left(1 - \frac{1}{2^{-H}}\right)^{n-3} P_{\text{err}}(k=2, \text{lh}=1) + \\
& + (n-3) \left(\frac{1}{2^{-H}}\right) \left(1 - \frac{1}{2^{-H}}\right)^{n-4} P_{\text{err}}(k=3, \text{lh}=1).
\end{aligned} \quad (11)$$

В общем виде для произвольных  $k$  и  $\text{lh}$  выражение (11) запишется как рекуррентное:

$$\begin{aligned}
P_{\text{err}}(k, \text{lh}) = & \sum_{i=2}^{n-2-k} C_{n-2-k}^i \left(\frac{1}{2^{-H} - \text{lh}}\right)^i \left(1 - \frac{1}{2^{-H} - \text{lh}}\right)^{n-2-k-i} + \\
& + \left(1 - \frac{1}{2^{-H} - \text{lh}}\right)^{n-2-k} P_{\text{err}}(k+1, \text{lh}+1) + \\
& + (n-2-k) \left(\frac{1}{2^{-H} - \text{lh}}\right) \left(1 - \frac{1}{2^{-H} - \text{lh}}\right)^{n-2-k-1} P_{\text{err}}(k+2, \text{lh}+1).
\end{aligned} \quad (12)$$

Итоговая вероятность построения  $P_{iz}$  изоморфных множеств ИП, для которых функция принадлежности будет истинной, а значения номеров пакетов в множестве вычислены в соответствии с формулами (4), будет равняться значению функции  $P_{\text{err}}(k, \text{lh})$  при  $k=2$  и  $\text{lh}=0$ . График зависимости вероятности возникновения ошибки построения структурированного множества от длины  $H$  поля  $s^{\text{hash}}$  и количества пакетов, на которые фрагментировано сообщение  $n$ , показан на рис. 2.

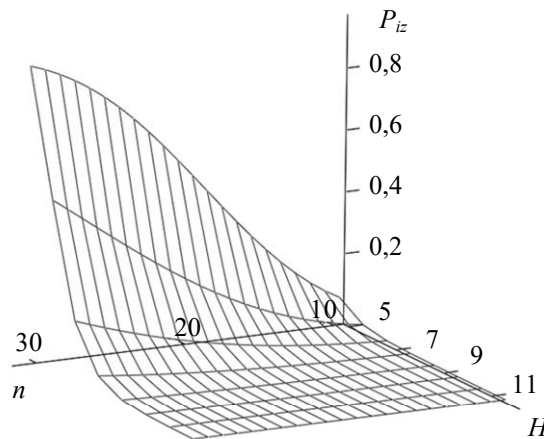


Рис. 2

Анализ графика показывает, что длина  $H$  поля  $s^{\text{hash}}$ , содержащего значение хеш-функции, оказывает существенное влияние на вероятность ошибки лишь до значения 6. Дальнейшее увеличение  $H$  снижает вероятность в пределах 1 % на каждый дополнительный бит длины. Также следует отметить, что значительный рост вероятности ошибки при небольших размерах поля, содержащего значение хеш-функции, происходит на интервале 10...30 пакетов, а затем наступает зона насыщения. При больших значениях  $H$  рост вероятности ошибки практически незначителен и не превышает 5 % к увеличению длины сообщения на каждые 10 пакетов.

На основании найденных зависимостей между вероятностью ошибки и размерами сообщения и длиной поля  $s^{\text{hash}}$  можно построить зависимость разрядности результата хеш-функции, которая обеспечивала бы заданную величину априорной ошибки, от размера сообщения в пакетах. Данная зависимость, приведенная на рис. 3, еще раз иллюстрирует сделан-

ный выше вывод, что увеличение длины  $H$  поля  $s^{\text{hash}}$  свыше 10 нецелесообразно, так как не дает значимого снижения вероятности ошибки построения сообщения.

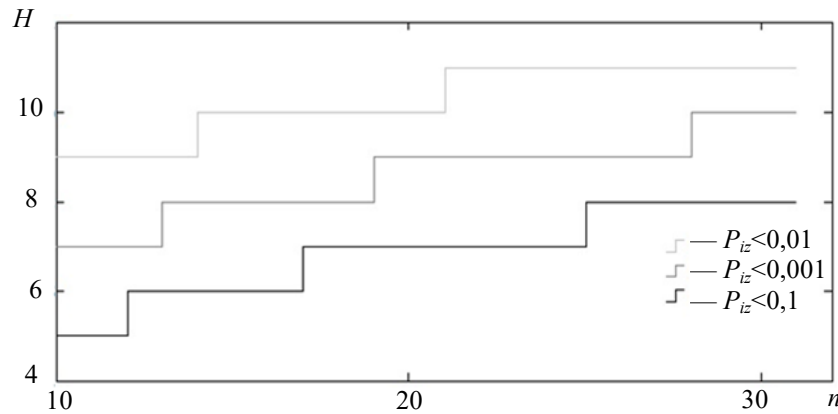


Рис. 3

Для определения оптимальной длины поля  $s^{\text{hash}}$  введем специальную меру — коэффициент использования канала передачи данных  $K$  как отношение объема информации, переданной в сообщении, к объему всех образующих его ИП, включая объем дополнительных полей и дополнительные стартовые и конечные ИП. При расчете данного параметра учитывалось, что возникшая ошибка построения требует повторной передачи данных. Число дополнительных пакетов, которые потребовались для повторной передачи сообщений, переданных с ошибками, при общем достаточно большом количестве  $N$  сообщений, определяется как сумма бесконечной убывающей геометрической прогрессии, показателем которой является вероятность возникновения ошибки определения функции принадлежности и порядкового номера каждого пакета [12]. Итоговая формула для коэффициента использования канала имеет следующий вид:

$$K = \frac{nL(1 - P_{iz})}{(L + H)(n + 2)}, \tag{13}$$

где  $H$  — длина поля, содержащего значение хеш-функции, полученное из данных предыдущего пакета;  $L$  — длина информационного поля пакета;  $P_{iz}$  — вероятность ошибки при заданной длине поля, содержащего значение хеш-функции.

График зависимости коэффициента  $K$  от длины сообщения  $n$  и длины информационного поля  $L$  при различных значениях  $H$  представлен на рис. 4 (здесь 1 —  $H=8$ , 2 —  $H=9$ ).

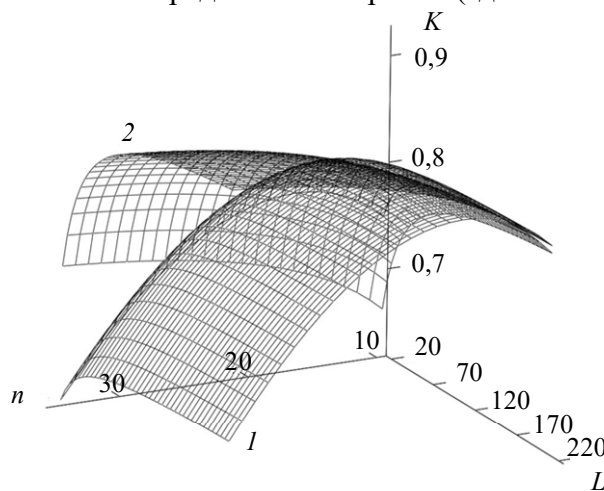


Рис. 4

Анализ графика показывает, что при фиксированном  $n$  с увеличением  $L$  коэффициент  $K$  асимптотически приближается к величине  $(1 - P_{iz})n(n + 2)^{-1}$ , что очевидным образом следует из формулы (13). Если рассматривать изменение данного коэффициента в зависимости от длины сообщения  $n$  при фиксированном параметре  $L$ , то график имеет характерный максимум, сначала повышаясь до него при росте  $n$ , а затем снижаясь. Интерпретацией данного факта может быть то, что при малой длине сообщения значительное влияние на снижение коэффициента использования канала оказывает необходимость передачи дополнительных стартового и конечного пакетов. С ростом числа ИП в сообщении дополнительная информационная избыточность, связанная с ними, уменьшается, а влияние возрастающей вероятности ошибки увеличивается. Данный максимум наблюдается до значений  $H < 10$ . При больших значениях  $H$  сама функция коэффициента использования канала передачи данных вырождается в асимптотически приближающуюся к значению  $L(L + H)^{-1}$  кривую.

Длина сообщения, при котором функция  $K$  достигает максимума, составляет некоторую область, слабо изменяющуюся с изменением параметра  $L$ , но находящуюся в зависимости от длины  $H$ . График данной зависимости, представленный на рис. 5, приведен до значений  $H = 10$ , так как при дальнейшем увеличении параметра  $H$  рост величины  $n_{\max}$  носит экспоненциальный характер. Практическое применение найденной зависимости заключается в том, что при наличии системных требований, которые ограничивают разрядность результата хеш-функции, например, из-за общего ограничения размера передаваемого ИП, можно подобрать длину единичного фрагментируемого сообщения  $n$  таким образом, чтобы достичь минимальной информационной избыточности, что скажется соответствующим образом на пропускной способности канала в целом и продолжительности автономной работы устройств, реализующих исследуемый алгоритм.

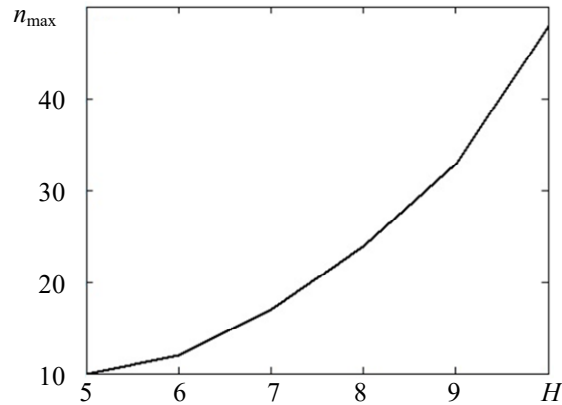


Рис. 5

Таким образом, используя представленный в настоящей статье вероятностный подход к определению источника информационного пакета и его порядкового номера во фрагментированном сообщении, можно добиться приемлемой для рассматриваемых протоколов достоверности.

#### СПИСОК ЛИТЕРАТУРЫ

1. Предварительный национальный стандарт РФ. ПНСТ 354-2019. Информационные технологии. Интернет вещей. Протокол беспроводной передачи данных на основе узкополосной модуляции радиосигнала (NB-Fi) [Электронный ресурс]: <<http://docs.cntd.ru/document/1200162760>>, 15.01.2020.
2. Предварительный национальный стандарт РФ. Информационные технологии. Интернет вещей. Протокол обмена для высокочастотных сетей с большим радиусом действия и низким энергопотреблением [Электронный ресурс]: <[https://drive.google.com/uc?id=12kPw5\\_ndO8zav7\\_BP\\_EXKdytu7uEyy3x&export=download](https://drive.google.com/uc?id=12kPw5_ndO8zav7_BP_EXKdytu7uEyy3x&export=download)>, 15.01.2020.

3. 802.15.4-2015 – IEEE Standard for Low-Rate Wireless Personal Area Networks // IEEE Computer Society. DOI:10.1109/ieeestd.2016.7460875.
4. Лоднева О. Н., Ромасевич Е. П. Анализ трафика устройств интернета вещей // Современные информационные технологии и ИТ-образование. 2018. Т. 14, № 1. С. 149—169.
5. Зайцев В., Соколов Н. Особенности мультисервисного трафика с учетом сообщений, создаваемых устройствами IoT // Первая миля. 2017. № 4. С. 44—47.
6. Papadimitratos P., Haas Z. Secure message transmission in mobile ad hoc networks // Ad Hoc Networks. 2003. № 1. P. 193—209.
7. Ben Othman S., Alzaid H., Trad A., Youssef H. An efficient secure data aggregation scheme for wireless sensor networks // PISA. 2013. DOI: 10.1109/iisa.2013.6623701.
8. Hayouni H., Hamdi M., Kim T. A novel efficient approach for protecting integrity of data aggregation in wireless sensor networks // Wireless Communications and Mobile Computing Conference (IWCMC). 2015. P. 1193—1198.
9. Wang W., Wang D., Jiang Y. Energy efficient distributed compressed data gathering for sensor networks. // Ad Hoc Networks. 2016. DOI: 10.1016/j.adhoc.2016.10.003.
10. Stallings W. NIST block cipher modes of operation for authentication and combined confidentiality and authentication // Cryptologia. 2010. N 34. P. 225—235.
11. Black J., Rogaway P. CBC MACs for arbitrary-length messages: The three-key constructions // Cryptologia. 2015. Vol. 18, N 2. P. 111—131.
12. Таныгин М. О., Алиаиа Х. Я., Алтухова В. А. Об одном методе контроля целостности передаваемой поблоково информации // Телекоммуникации. 2019. № 3. С. 12—21.
13. Арсеньев В. Н., Силантьев С. Б., Ядренкин А. А. Использование априорной информации для коррекции модели потока событий в сложной системе // Изв. вузов. Приборостроение. 2017. Т. 60, № 5. С. 391—397.
14. Tanygin M. O., Tipikin A. P. Methods of authentication of information protection systems and controlling software // Telecommunications and Radio Engineering. 2007. Vol. 66, N 5. P. 453—463.
15. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука, 1978. 832 с.

#### Сведения об авторе

**Максим Олегович Таныгин**

— канд. техн. наук, доцент; Юго-Западный государственный университет, кафедра информационной безопасности; заведующий кафедрой; E-mail: tanygin@yandex.ru

Поступила в редакцию  
28.02.2020 г.

**Ссылка для цитирования:** Таныгин М. О. Алгоритм определения источника фрагментированных сообщений // Изв. вузов. Приборостроение. 2020. Т. 63, № 8. С. 702—710.

## ALGORITHM FOR DETERMINING THE SOURCE OF FRAGMENTED MESSAGES

**M. O. Tanygin**

*South-West State University, 305040, Kursk, Russia  
E-mail: tanygin@yandex.ru*

The problem of reducing information redundancy when exchanging data in networks with a long range and low power consumption is analyzed. The approach under consideration is based on reduction of the size of additional service fields that determines the information packet source and its number in the single message. To increase the entropy of these additional service fields, it is proposed to write in them a hash formed from data from other information packets of the source and the source unique identifier. A formalized algorithm for analyzing incoming information packets by the receiver which allows determining the source of the information packet is described. The problem of an error in determining the sequence number of a fragment in a single message is formulated. Based on the mathematical apparatus of the probability theory, numerical characteristics of such type of error are derived, and its influence on the information redundancy is estimated. Dependencies between the additional hash field length and number of



the transmitted message fragments are presented, and the range of the values providing the minimum information redundancy is determined.

**Keywords:** messages source determination, probability theory, hash, packet order violation, algorithm parameters, information redundancy

#### REFERENCES

1. <http://docs.cntd.ru/document/1200162760>. (in Russ.)
2. [https://drive.google.com/uc?id=12kPw5\\_ndO8zav7\\_BP\\_EXKdytu7uEyy3x&export=download](https://drive.google.com/uc?id=12kPw5_ndO8zav7_BP_EXKdytu7uEyy3x&export=download). (in Russ.)
3. 802.15.4-2015 – IEEE Standard for Low-Rate Wireless Personal Area Networks, IEEE Computer Society. DOI:10.1109/ieeestd.2016.7460875.
4. Lodneva O.N., Romasevich E.P. *Modern Information Technologies and IT-education*, 2018, no. 1(14), pp. 149–169. (in Russ.)
5. Zaytsev V., Sokolov N. *Pervaya milya*, 2017, no. 4, pp. 44–47. (in Russ.)
6. Papadimitratos P., Haas Z.J. *Ad Hoc Networks*, 2003, no. 1, pp. 193–209.
7. Ben Othman S., Alzaid H., Trad A., Youssef H. *IISA*, 2013. DOI:10.1109/iisa.2013.6623701.
8. Hayouni H., Hamdi M., Kim T. *Wireless Communications and Mobile Computing Conference (IWCMC)*, 2015, pp. 1193–1198.
9. Wang W., Wang D., Jiang Y. *Ad Hoc Networks*, 2016. DOI: 10.1016/j.adhoc.2016.10.003.
10. Stallings W. *Cryptologia*, 2010, no. 34, pp. 225–235.
11. Black J., Rogaway P. *Cryptologia*, 2015, no. 2(18), pp. 111–131.
12. Tanygin M.O., Alshaia Kh.Ya., Altukhova V.A. *Telekommunikatsii* (Telecommunications), 2019, no. 3, pp. 12–21. (in Russ.)
13. Arsen'yev V.N., Silant'yev S.B., Yadrenkin A.A. *Journal of Instrument Engineering*, 2017, no. 5(60), pp. 391–397.
14. Tanygin M.O., Tipikin A.P. *Telecommunications and Radio Engineering*, 2007, no. 5(66), pp. 453–463.
15. Korn G.A., Korn T.M. *Mathematical handbook: For scientists and engineers*, NY, McGraw-Hill, 1968.

#### Data on author

**Maxim O. Tanygin** — PhD, Associate Professor; South-West State University, Department of Information Security; Head of the Department; E-mail: tanygin@yandex.ru

**For citation:** Tanygin M. O. Algorithm for determining the source of fragmented messages. *Journal of Instrument Engineering*. 2020. Vol. 63, N 8. P. 702–710 (in Russian).

DOI: 10.17586/0021-3454-2020-63-8-702-710