

Н.А. Штомпель, канд. техн. наук, доцент  
Український державний університет залізничного транспорту, г. Харків, Україна  
shtompel.mykola@kart.edu.ua

## Биоинспирированный метод оптимизации кодов на основе преобразования Лаби

*Обосновано применение кодов на основе преобразования Лаби для обеспечения заданного качества обслуживания в компьютерных сетях. Представлены особенности процессов кодирования и декодирования информации кодами Лаби. Показано, что эффективность кодов на основе преобразования Лаби значительным образом зависит от закона распределения вероятностей, на основе которого формируются кодовые символы. Предложен подход к оптимизации относительно длинных кодов Лаби с использованием обобщенных биоинспирированных процедур. Рассмотрены особенности и основные этапы предложенного метода оптимизации кодов на основе преобразования Лаби.*

**Ключевые слова:** компьютерные сети, коды Лаби, оптимизация, биоинспирированные процедуры.

### Введение

В настоящее время для передачи разнообразной информации (речь, видео, данные) широко используются компьютерные сети, которые строятся на основе коммутации пакетов. При этом проблема потери пакетов значительным образом влияет на качество обслуживания пользователей данных сетей [1]. Классическим подходом к восстановлению потерянных пакетов в компьютерных сетях на основе стека протоколов TCP/IP является применение методов повторной передачи, которые лежат в основе протокола TCP. Однако данный подход можно использовать только при передаче данных, т.к. в этом случае допускается значительная величина задержки пакетов. С другой стороны, с учетом значительного роста числа приложений, работающих в реальном масштабе времени и использующих «ненадежный» протокол UDP, возникает задача обеспечения заданного качества обслуживания путем применения других методов восстановления потерянных пакетов [2].

### Постановка проблемы и анализ литературы

Современным подходом к восстановлению потерянных пакетов в компьютерных сетях является применение фонтанных кодов, входящих в состав разнообразных протоколов. В данном случае исходный блок данных разделяется на несколько частей некоторой длины – информационные символы. Отправитель по определенному правилу формирует кодовые пакеты (кодовые символы). Процесс кодирования теоретически может быть бесконечным и независимым для отдельных информационных символов. Таким образом, фонтанные коды не имеют фиксированной скорости кодирования в

отличии от классических кодовых конструкций. Получатель может декодировать принятые символы после накопления заданного числа произвольных кодовых пакетов [3].

Коды на основе преобразования Лаби (коды Лаби) являются первыми фонтанными кодами, которые стали применяться на практике [4]. Отличительная особенность данных кодов заключается в использовании специального закона (плотности) распределения вероятностей, на основе которого формируются кодовые символы. При этом эффективность кодов Лаби существенно зависит от данного распределения. В [5] обосновано применение робастного солитонного распределения вероятностей для длинных кодов Лаби. С другой стороны, многие современные прикладные приложения ориентированы на использование относительно коротких кодов Лаби. В [6], [7] предложены улучшенные солитонноподобные законы распределения вероятностей для очень коротких кодов Лаби. В [8] представлен подход к оптимизации законов распределения вероятностей для коротких кодов Лаби на основе эволюционной стратегии с адаптацией ковариационной матрицы. Известные методы формирования законов распределения вероятностей для кодов Лаби имеют ряд недостатков, в частности, характеризуются высокой вычислительной сложностью и возможностью применения только для кодов малой длины.

Таким образом, актуальной задачей является обеспечение заданного качества обслуживания в компьютерных сетях в условиях потери пакетов путем оптимизации относительно длинных кодов на основе преобразования Лаби.

Целью статьи является повышение эффективности кодов на основе преобразования

Лабн для обеспечения заданного качества обслуживания в компьютерных сетях.

### Основная часть

Кодирование информации кодами Лабн предполагает разделение исходного сообщения на  $k$  информационных символов, состоящих из  $l$  бит. Кодовый символ формируется путем сложения по модулю 2 нескольких информационных символов. Графически процедуру кодирования можно представить в виде двудольного графа, состоящего из информационных и кодовых вершин. Значительное влияние на эффективность кодов на основе преобразования Лабн оказывает распределение степеней кодовых вершин  $\rho(d)$  в данном графе, которое удовлетворяет следующим условиям  $\sum_{d=1}^k \rho(d) = 1$  и  $1 \leq d \leq k$ .

Фактически  $\rho(d)$  представляет собой плотность распределения вероятностей случайной величины  $d$ , которая определяет количество информационных символов, используемых для формирования некоторого кодового символа. Данные информационные символы выбираются случайным образом из  $k$  информационных символов и называются «соседями».

Декодирование кодов Лабн может осуществляться с использованием нескольких подходов [3], например, путем применения метода декодирования на основе распространения доверия. Данный метод декодирования характеризуется невысокой вычислительной сложностью, но требует избыточного количества принятых кодовых символов  $n$  ( $n > k$ ), в отличие от метода декодирования на основе исключения Гаусса. Рассмотрим особенности процесса декодирования кодов Лабн для модели стирающего канала связи, которая может использоваться для формального представления передачи пакетов в компьютерных сетях [9]. При инициализации декодирования все принятые информационные символы считаются неизвестными (невосстановленными). На первом этапе кодовые символы, имеющие только одного «соседа» могут быть непосредственно использованы для однозначного определения соответствующих информационных символов. В дальнейшем данные кодовые символы исключаются из процесса декодирования. Если информационный символ может быть восстановлен, но еще не был обработан, он называется «ripple»-символом и помещается в очередь. На каждом последующем этапе декодирования «ripple»-символы последовательно извлекаются из очереди и удаляются из кодовых символов, которые являются их «соседями». Если после данной процедуры кодовый символ

содержит только одного «соседа», то осуществляется восстановление соответствующего информационного символа. В дальнейшем могут быть сформированы новые «ripple»-символы для сохранения стабильного размера очереди. Поддержание заданного размера очереди обеспечивает продолжение процесса декодирования до восстановления всех информационных символов. В противном случае, если «ripple»-символы будут отсутствовать, то произойдет отказ от декодирования или понадобится продолжить прием кодовых символов.

Следует отметить, что эффективность декодирования кодов Лабн на основе распространения доверия требует использования «хорошего» распределения степеней кодовых вершин  $\rho(d)$ . При этом с практической точки зрения целесообразно обеспечить наименьшую избыточность, которая равна:

$$\varepsilon = \frac{n}{k}. \quad (1)$$

В [5] предложено идеальное солитонное распределение:

$$\sigma(d) = \begin{cases} \frac{1}{k}, & d=1, \\ \frac{1}{d(d-1)}, & d=2, \dots, k. \end{cases} \quad (2)$$

Распределение (2) гарантирует, что на каждой итерации декодирования будет освобождаться кодовый символ с вероятностью  $1/k$ . Таким образом, на каждом этапе декодирования для  $k$  информационных символов ожидается только один «ripple»-символ. В идеальном случае, все информационные символы будут восстановлены после  $k$  итераций. Следовательно, использование идеального солитонного распределения теоретически позволяет получить коды Лабн с избыточностью  $\varepsilon = 1$ .

Однако, данное распределение не учитывает случайный характер процессов кодирования и декодирования, что приводит к ошибкам декодирования из-за отсутствия «ripple»-символа на некотором этапе декодирования. В соответствии с теорией случайных блужданий, вероятность с которой случайное блуждание длиной  $k$  отклоняется от своего среднего значения более чем на  $\ln(k/\delta)\sqrt{k}$  не превышает  $\delta$ . Данный факт лежит в основе выбора размера очереди из «ripple»-символов, который гарантирует успешное завершение процесса декодирования.

Представленный выше подход позволяет реализовать робастное солитонное распределение:

$$\mu(d) = \frac{\sigma(d) + \tau(d)}{\beta}, \quad (3)$$

где  $\tau(d)$  – функция усиления, которая определяется по формуле:

$$\tau(d) = \begin{cases} \frac{R}{dk}, & d = 1, 2, \dots, (k/R) - 1, \\ \frac{R}{k} \ln(R/\delta), & d = k/R, \\ 0, & d > k/R, \end{cases}$$

где  $\delta$  – параметр, определяющий вероятность успешного декодирования;

$R$  – вспомогательная величина, рассчитываемая по формуле:

$$R = c \cdot \ln(k/\delta) \cdot \sqrt{k},$$

где  $c$  – параметр, определяющий среднее значение распределения степеней кодовых вершин;

$\beta$  – нормирующий множитель, который равен:

$$\beta = \sum_{d=1}^k \sigma(d) + \tau(d).$$

Изменение характеристик распределения (3) осуществляется с помощью параметров  $\delta$  и  $c$ . Параметр  $\delta$  позволяет изменять размер очереди из «ripple»-символов, определяемый как  $\ln(k/\delta)\sqrt{k}$ . Уменьшение значения параметра  $c$  увеличивает вероятность низких степеней кодовых вершин, что снижает вычислительную сложность процесса кодирования. При использовании данного распределения для успешного восстановления информационных символов с вероятностью не менее  $1 - \delta$ , требуется порядка  $n = k + O(\ln^2(k/\delta)\sqrt{k})$  кодовых символов.

Асимптотический анализ робастного солитонного распределения показал, что оно является близким к оптимальному [5]. Однако, в реальных компьютерных сетях исходное сообщение не может быть разбито на блоки данных бесконечной длины, т.е. число информационных символов всегда является конечной величиной. В результате этого эффективность кодов Лаби с робастным солитонным распределением существенно снижается при уменьшении количества информационных символов  $k$ .

Для устранения данного ограничения предлагается развить подход [8] к оптимизации распределения степеней кодовых вершин кодов Лаби с использованием обобщенных биоинспирированных процедур.

Рассмотрим формальное представление задачи поиска распределения степеней кодовых вершин  $\rho(d)$  кода Лаби с заданным числом информационных символов  $k$  в виде оптимизационной задачи.

Как было показано выше, «хорошее» распределение степеней кодовых вершин должно обеспечивать минимальную избыточность (1), величину которой целесообразно принять в качестве целевой функции. Искомую плотность распределения вероятностей можно представить в виде вектора длиной  $k$ , элементами которого являются нормализованные вещественные числа. Для уменьшения размерности оптимизационной задачи для относительно длинных кодов Лаби целесообразно заранее выбрать позиции из множества  $D$  с ненулевыми значениями вероятностей в исходном векторе. Из анализа особенностей солитонных распределений и декодирования на основе распространения доверия следует, что искомое распределение степеней кодовых вершин обязательно должно иметь ненулевую вероятность кодовой вершины степени 1. Следовательно, оптимизируемое распределение степеней кодовых вершин можно представить следующим образом:

$$\omega(d) = \begin{cases} \omega_1, & d = 1, \\ \omega_d, & d \in D, \\ 0, & d \notin D. \end{cases} \quad (4)$$

Поиск оптимального распределения (4) для заданного значения  $k$  предлагается осуществлять путем компьютерного моделирования передачи информации через стирающий канал с использованием кодов Лаби. Данные коды имеют случайный характер процессов кодирования и декодирования, что не позволяет гарантировать восстановление всей информации на основе фиксированного числа принятых кодовых символов. Поэтому для того, чтобы оценить избыточность  $\varepsilon$  генерируется непрерывная последовательность кодовых символов, которые декодируются до момента, пока не будет восстановлено исходное информационное сообщение. Таким образом, в данном случае среднее число требуемых кодовых символов, получаемых в результате моделирования, выступает в роли целевой функции.

В качестве поискового механизма в данной работе предлагается использовать обобщенные биоинспирированные процедуры, которые включают в себя эволюционные процедуры; популяционные процедуры, вдохновленные живой и неживой природой; процедуры, вдохновленные человеческим обществом и т.д. При этом конкретные биоинспирированные процедуры характеризуются различной вычислительной сложностью [10]. Таким образом, эволюционная стратегия с адаптацией ковариационной матрицы, используемая в [8] для оптимизации кодов Лаби, является частным случаем предлагаемого подхода.

Общая схема популяционных биоинспирированных процедур содержит

следующие этапы: инициализация популяции, миграция агентов популяции, окончание поиска.

С учетом особенностей распределения (4), в качестве агентов популяции целесообразно использовать векторы  $\varpi = (\omega_1, \omega_d)$ , которые содержат только ненулевые элементы данного распределения. При инициализации популяции генерируется заданное число векторов  $\varpi$ , значения элементов которых подчиняются равномерному распределению. Далее полученные векторы преобразуются в соответствующие распределения степеней кодовых вершин  $\omega(d)$  для последующего кодирования информации кодом Лаби с определенным параметром  $k$ . После этого осуществляется моделирование процесса передачи информации по каналу связи и восстановление переданного сообщения путем декодирования на основе распространения доверия принятых кодовых символов. По результатам моделирования вычисляется среднее значение избыточности для заданного кода Лаби, которое является фитнес-функцией для обобщенных биоинспирированных процедур. Затем происходит перемещение агентов популяции в области поиска с использованием некоторого набора миграционных операторов, т.е. фактически осуществляется изменение векторов  $\varpi$ , для получения наименьшего значения фитнес-функции. Окончание вычислений происходит при достижении заданного числа итераций обобщенных биоинспирированных процедур, а в качестве приближенного решения выбирается наилучшее положение агентов популяции – вектор  $\varpi^*$ .

Таким образом, поиск «хорошего» распределения степеней кодовых вершин кода Лаби для заданной величины числа информационных символов формально можно представить в виде задачи нелинейного программирования:

$$\begin{aligned} f(\varpi^*) &= \min_{\varpi \in \varpi'} f(\varpi), \\ f(\varpi) &= \varepsilon, \\ k &= \text{const}, \\ \varpi' &= \left\{ \varpi \left\{ \begin{array}{l} 0 < \omega_1 < 1, \\ 0 < \omega_d < 1 \\ \sum_{d=1}^k \omega(d) = 1 \\ 1 \leq d \leq k \\ d = 1, d \in D \end{array} \right. \right\}, \end{aligned}$$

где  $\varpi'$  – множество допустимых решений, соответствующее группе векторов  $\varpi$ .

### Выводы

В современных компьютерных сетях для восстановления потерянных пакетов целесообразно применять коды Лаби. Предложенный метод оптимизации на основе обобщенных биоинспирированных процедур позволяет повысить эффективность относительно длинных кодов Лаби. В дальнейшем планируется оценить вычислительную сложность оптимизации кодов Лаби с использованием разных биоинспирированных процедур.

### Список литературы

1. Олифер В.Г. Компьютерные сети. Принципы, технологии, протоколы: учебник для вузов. 4-е изд. / В.Г. Олифер, Н.А. Олифер. – СПб.: Питер, 2010. – 944 с.: ил.
2. Асауленко І.О. Аналіз методів відновлення даних у телекомунікаційних мережах з комутацією пакетів / І.О. Асауленко, С.І. Приходько, М.А. Штомпель // Інформаційно-керуючі системи на залізничному транспорті: науково-технічний журнал. – Харків: УкрДУЗТ, 2015. – Вип. 4 (113). – С. 27 – 38.
3. Варгаузин В. Помехоустойчивое кодирование в пакетных сетях / В. Варгаузин // ТелеМультиМедиа. – 2005. – №3. – С. 10 – 16.
4. Шинкаренко К.В. Помехоустойчивое кодирование мультимедиа данных в компьютерных сетях / К.В. Шинкаренко, А.М. Кориков // Известия Томского политехнического университета. – 2008. – Т. 313, №5. – С. 37– 41.
5. Luby M. LT codes // Proceedings of the 43rd Symposium on Foundations of Computer Science, IEEE Computer Society, 2002. – pp. 271 – 280.
6. Bodine E.A., Cheng M. K. Characterization of Luby transform codes with small message size for low-latency decoding // IEEE International Conference on Communications (ICC '08), 2008. – pp. 1195 – 1199.
7. Nyutia E., Tirronen T., Virtamo J. Optimal degree distribution for LT codes with small message length // Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM 2007), 2007. – pp. 2576 – 2580.
8. Chen C.M., Chen Y.P., Shen T.C., Zao, J.K. On the optimization of degree distributions in LT code with covariance matrix adaptation evolution strategy // IEEE Congress on Evolutionary Computation, 2010, July. – pp. 1 – 8.

9. Асауленко І.О. Дослідження характеристик телекомунікаційних систем з використанням програмних реалізацій каналів зв'язку / І.О. Асауленко, М.А. Штомпель // Інформаційно-керуючі системи на залізничному транспорті. – 2015. – № 1. – С. 34 – 41.

10. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учебное пособие / А.П. Карпенко. – М: издательство МГТУ им. Н.Э. Баумана, 2014. – 446 с.

*Надійшла до редакції 20.10.2016*

### **М.А. ШТОМПЕЛЬ**

Український державний університет залізничного транспорту

#### **БІОІНСПІРОВАНІЙ МЕТОД ОПТИМІЗАЦІЇ КОДІВ НА ОСНОВІ ПЕРЕТВОРЕННЯ ЛАБИ**

Обґрунтовано застосування кодів на основі перетворення Лабі для забезпечення заданої якості обслуговування в комп'ютерних мережах. Представлені особливості процесів кодування і декодування інформації кодами Лабі. Показано, що ефективність кодів на основі перетворення Лабі значним чином залежить від закону розподілу ймовірностей, на основі якого формуються кодові символи. Запропоновано підхід до оптимізації відносно довгих кодів Лабі з використанням узагальнених біоінспірованих процедур. Розглянуто особливості та основні етапи запропонованого методу оптимізації кодів на основі перетворення Лабі.

**Ключові слова:** *комп'ютерні мережі, коди Лабі, оптимізація, біоінспіровані процедури.*

### **М.А. SHTOMPEL**

Ukrainian State University of Railway Transport

#### **BIOSPIRED OPTIMIZATION METHOD CODES BASED ON LUBY TRANSFORM**

It has been shown that the packet loss problem considerably affects the quality of service users of computer networks. The limitations of classical approaches to the recovery of the lost packets are considered. The principles of encoding and decoding fountain codes are given. It has been shown that the feature codes based on Luby transform is the using of a special law (density) of probability distribution for forming code symbols. The features soliton-like laws of probability distribution for Luby codes are considered. It has been shown that the efficiency of Luby codes with robust soliton probability distribution is significantly reduced with decreasing the number of information symbols. The optimization approaches for codes based on Luby transform different lengths are analyzed. The necessity of optimizing the laws of probability distribution for a relatively long Luby codes is justified. The approach to optimization of these codes with using generalized bioinspired procedures is proposed. The general scheme of population bioinspired procedures of search optimization is provided. The features and main stages of the proposed optimization method codes based on Luby transform are considered. During the first stage of the proposed approach to optimization Luby codes generates a set of vectors, which corresponding to different probability distributions. Further encoding information messages with Luby codes, modeling the information transmission over a communication channel and the recovery of the transmitted messages by decoding based on the belief propagation received code symbols are performed. Then search the best probability distribution for specified parameters Luby code with the help of generalized bioinspired procedures is carried out. The formal presentation of search "good" probability distribution for Luby codes in the form of non-linear programming problem is considered. It has been shown that as the objective function is advisable to take the amount of Luby codes redundancy. The restrictions for the formulated problem of minimizing objective function are provided.

**Keywords:** *computer networks, Luby codes, optimization, bioinspired procedures.*