DOI: 10.25728/mlsd.2022.1231

ПОДХОДЫ К ПРЕОБРАЗОВАНИЮ БОЛЬШИХ МАССИВОВ ИНФОРМАЦИИ В ЦЕПЬ БЛОКОВ ЦИФРОВЫХ ДАННЫХ, ЗАЩИЩЕННУЮ ОТ МОДИФИКАЦИИ ЧЛБ (N, K)-КОДОМ

Терентьев А.И.

Московский государственный технический университет гражданской авиации (МГТУ ГА)
Россия, г. Москва, Кронштадтский бульвар, д. 20
terentyev-doc@yandex.ru

Аннотация: Предложены некоторые подходы к преобразованию больших массивов информации в линейно упорядоченные множества (цепи) блоков цифровых данных, защищенные от модификации посредством кодирования их элементов (блоков) числовым линейным блоковым корректирующим кодом (ЧЛБ-кодом) над кольцом конечных десятичных дробей. Введено название таких множеств – ЧЛБ-цепи. Рассмотрены отдельные аспекты построения ЧЛБ-цепей на практике.

Ключевые слова: числовой линейный блоковый корректирующий код (ЧЛБ-код), технология связных данных, ЧЛБ-цепь, технология блокчейн.

Введение

В эпоху цифровой трансформации особенно актуальной является задача преобразования различных массивов информации, независимо от их отраслевой принадлежности, в вид, адекватный для процессов и технических систем, использующих такую информацию. При этом, в зависимости от назначения, функциональных, конструктивных, эксплуатационных, потребительских и иных особенностей конкретных технических систем, возможны различные требования, предъявляемые к структуре и свойствам, как самих массивов информации в целом, так и к составляющим их отдельным информационным единицам (объектам, элементам). Одним из таких требований может являться специальное структурирование (преобразование) исходного массива информации и представление его в виде линейно упорядоченного множества (цепи) блоков цифровых данных (далее используется название – цепь блоков цифровых данных), что представляется весьма удобным для последующей обработки и использования полученной цепи в технических системах, построенных на базе электронных вычислительных машин и устройств.

Учитывая современные угрозы любой ценной информации в цифровом пространстве, в том числе связанные с нарушением ее целостности и доступности, подавляющее большинство собственников информации и разработчиков использующих ее технических систем уже на этапе проектирования предусматривают защиту цепи блоков цифровых данных от преднамеренной и не преднамеренной (случайной) модификации. В настоящее время с целью обеспечению такой защиты применяются различные методы и технологии. В частности, благодаря созданию и развитию сети криптовалюты Биткоин (англ. Bitcoin) [1] широкую известность получила технология блокчейн (англ. Blockchain), позволяющая обеспечить построение защищенной от модификации цепи блоков цифровых данных. Однако, указанная технология имеет ряд существенных недостатков, что побуждает научное и профессиональное сообщество искать иные пути обеспечения защиты от модификации линейно упорядоченных множеств (цепей) блоков цифровых данных.

1 Общие подходы к преобразованию массивов информации в цепь блоков цифровых данных, защищенную от модификации

В качестве основы для подходов к преобразованию массивов информации в цепь блоков цифровых данных, защищенную от модификации, предлагается использовать концептуальную схему (парадигму) технологии связных данных [2], в которой любые данные, локализованные во времени и пространстве, рассматриваются с позиции математики как некоторое множество M, элементами которого могут являться множества, а также конструктивные, гибридные и иные сложно структурированные объекты, включающие, в том числе, информационную и служебную составляющие.

Согласно [бит] любая информация, независимо от ее содержательности, ценности и полезности, а также первоначальных формы и вида представления, в зависимости от поставленной практической задачи, может быть представлена как некоторые данные или их наборы. При этом в случае использования массива содержащаяся в нем информация, как правило, разбита на условные информационные единицы (объекты, элементы), определенным образом структурирована и упорядочена, что существенно упрощает задачу идентификации и индексирования (нумерации) таких

элементов, а также объединения их (при необходимости) в блоки. Таким образом, независимо от размера и вида массива информации, всегда существует возможность преобразования его в некоторое множество M, которое может быть задано посредством перечисления всех его элементов: $M = \{m_1, m_2, \ldots, m_k\}$, или посредством указания порождающей процедуры (алгоритма) или свойства, на основании которого они принадлежат этому множеству: $M = \{m|P(m)\}$.

Множество M будем считать связным, если на нем установлено бинарное или иное отношение R, обладающее свойством связности. Соответственно данные, которые представляются элементами такого множества будут являться в той или иной степени связанными между собой.

Порождающая процедура (свойство, функция, алгоритм), посредством которого устанавливается взаимосвязь между элементами множества и обеспечивается его связность, может быть любой природы, в зависимости от решаемой научной или практической задачи. В качестве примера в [2, 3] приведены следующие виды порождающих процедур:

- математической (числовой, логической, аналитической и т.п.);
- криптографической;
- семантической;
- визуально-смысловой;
- лексикографической;
- табличной;
- комбинированной и иной.

На связном множестве M может быть установлено отношение порядка, при котором нумерация (индексирование) его элементов будет не случайной (будет носить регулярный характер).

Отношение порядка может устанавливаться тем или иным способом, в том числе рекуррентным. Если отношение порядка установлено рекуррентным способом, то в простейшем (одномерном) случае множество M можно рассматривать как рекуррентную последовательность, в которой для любой пары элементов будет выполняться отношение строгого порядка $m_i R m_j$, а для каждой пары соседних элементов (m_{i-1}, m_i) или (m_i, m_{i+1}) рекуррентной последовательности будет выполняться отношение доминирования $m_{i-1} R m_i$ и $m_i R m_{i+1}$. Абстрактное графическое представление такой последовательности в ортогональном базисе приведено на рис. 1.

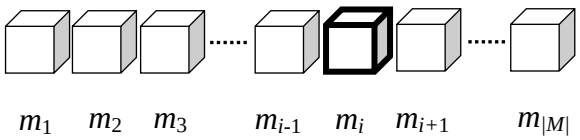

Рис. 1. Абстрактное графическое представление рекуррентной последовательности (n=1)

При этом в [2] показано, что в зависимости от решаемой практической задачи установление связности элементов и отношения порядка на множестве (совокупности множеств) может быть не только линейным, но и нелинейным, односвязным или многосвязным, а также n-мерным (многомерным), что в свою очередь будет определять структуру построенного множества M. В случае многомерности ($n \geq 2$) такое множество предложено называть системой связных множеств (подмножеств). В качестве примера одномерной ($n = 1$) структуры в [2] приведены последовательность (цепь) и кольцо, а двумерной ($n = 2$) – таблица или матрица (в ортогональном базисе), а также труба или объекты спирально-винтового типа. Простейшим примером трехмерной ($n = 3$) структуры множества M является куб или параллелепипед (в ортогональном базисе).

Любые данные могут быть тем или иным способом представлены в виде наборов чисел или одного числа. Если средством обработки таких данных является электронное техническое устройство (электронная вычислительная машина), то соответствующие числа или число будут являться элементами кольца конечных десятичных дробей D.

Таким образом, можно получить связное перечислимое разрешимое и линейно упорядоченное множество – цепь блоков цифровых данных, допускающее, при необходимости, дополнение его новыми элементами (блоками). Учитывая, что такая цепь будет являться объектом обработки в различных технических системах, ее важнейшим свойством в условиях современного цифрового пространства является способность противостоять преднамеренной и не преднамеренной модификации, что достигается посредством применения к блокам цепи специальных методов и процедур.

2 ЧЛБ-цепь и порождающая ее процедура на основе ЧЛБ (*n*, *k*)-кода

В [4, 5, 6] предлагается в качестве альтернативы технологии блокчейн использовать процедуру построения устойчивой к модификации цепи блоков цифровых данных, основанную на методах кодирования элементов (блоков) цепи числовым линейным блоковым разделимым систематическим корректирующим кодом (ЧЛБ (*n, k*)-кодом) над кольцом конечных десятичных дробей *D*. Элементы теории и практики ЧЛБ-кодов опубликованы в [7, 8].

Суть идеи, изложенной в [4, 5, 6], заключается в том, что элементами кодовых комбинаций ЧЛБ (*n, k*)-кода должны являться числовые представления блоков цифровых данных, а не составляющие эти блоки отдельные наборы цифровых данных. Следует отметить, что это не исключает, при необходимости, внутреннее кодирование каким-либо корректирующим кодом цифровых данных, входящих в блок. Параметры ЧЛБ (*n, k*)-кода, а именно количество информационных элементов его кодовых комбинаций, определяют число блоков цепи, подлежащих кодированию в конкретный момент времени. При этом последним, т.е. *k*-тым информационным элементом формируемой кодовой комбинации u_i, будет являться блок, добавляемый в цепь в соответствующий момент времени. При формировании и добавлении в следующий момент времени к цепи нового, т.е. $i + 1$, блока, нумерация информационных элементов в кодовой комбинаций u_{i+1} ЧЛБ (*n, k*)-кода сдвигается в сторону этого блока, который на текущем этапе будет являться *k*-тым информационным элементом этой кодовой комбинации. Графическая иллюстрация описанного метода показана на рис. 2, где пунктирными линиями обозначены блоки цепи, которые будут формироваться и включаться в цепь в последующие моменты времени.

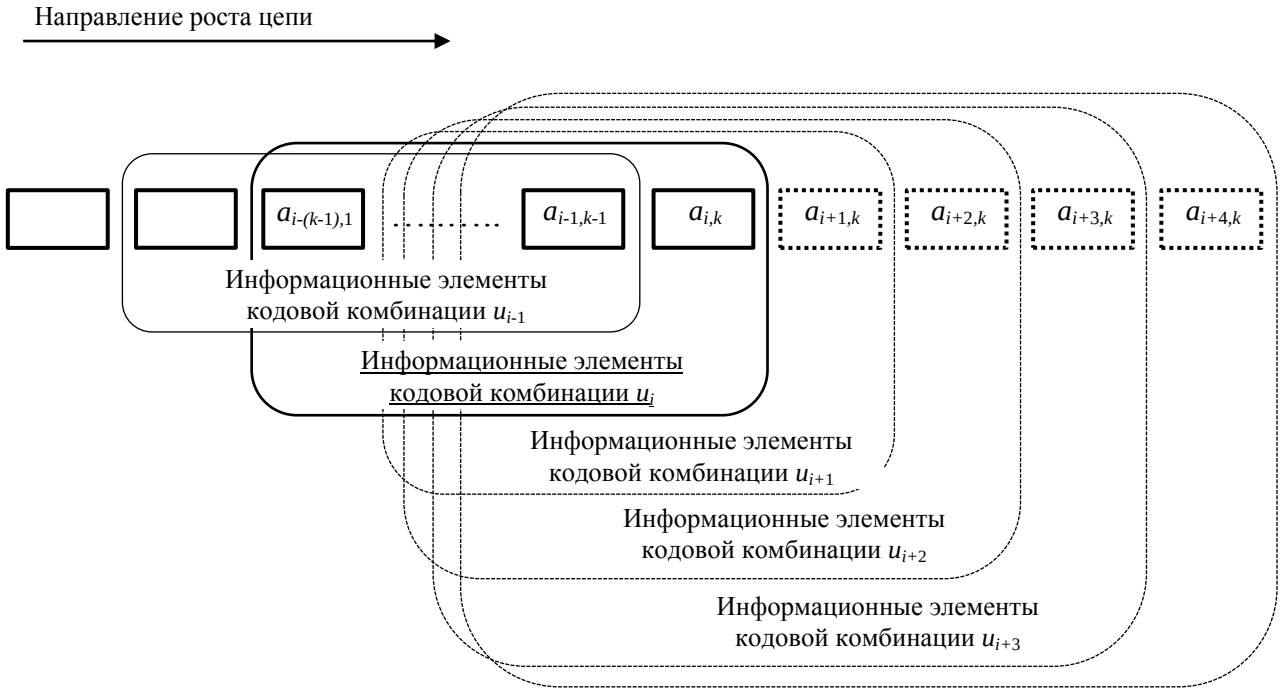

Рис. 2. Иллюстрация принципа формирования последовательностей информационных элементов кодовых комбинаций $u \in U_k$, где U_k – равномерный разделимый систематический ЧЛБ (n, k)-код над кольцом конечных десятичных дробей D

В целях упрощения визуального восприятия в [6] продемонстрирована иная графическая интерпретация принципа выбора блоков цифровых данных для последующего формирования из них последовательности информационных элементов, подлежащих кодированию достаточно коротким ЧЛБ (n_1, k_1)-кодом с $k = 4$. В качестве координатных осей выбраны «время» и «мощность множества $|M|$», т.е. количество блоков (длина) цепи. Также в целях усложнения задачи для потенциальных злоумышленников показан способ увеличения многомерности числового кодирования, а следовательно, и повышения результирующей силы установленной связи между блоками цифровых данных, за счет формирования и кодирования вторым ЧЛБ (n_2, k_2)-кодом последовательностей информационных элементов по дополнительному условному направлению (измерению), как это показано на рис. 3. При этом последовательности информационных элементов, которые выделены жирными контурами, соответствуют состоянию цепи блоков цифровых данных в моменты времени

t_1, …, t_6. Блоки цифровых данных, которые не используются в моменты времени t_1, …, t_6 для вычисления кодовых комбинаций, обозначены пунктирными контурами.

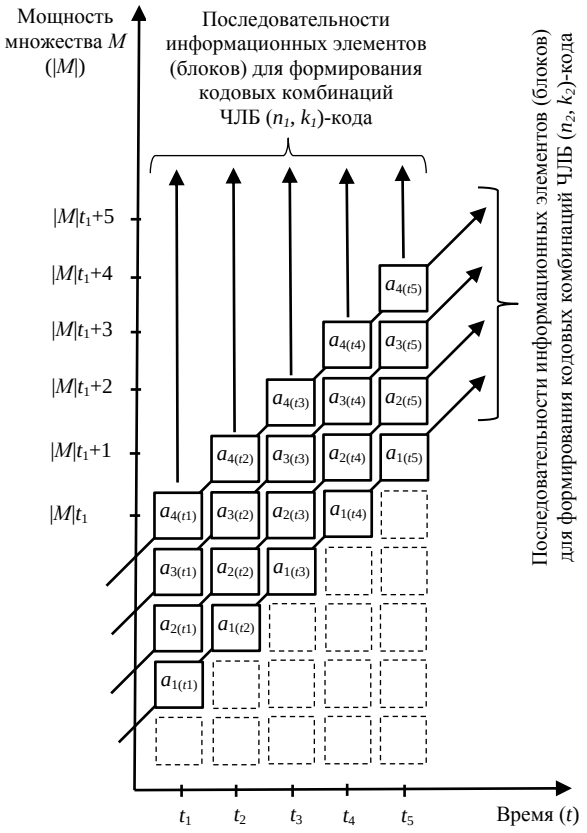

Рис. 3. Графическая иллюстрация формирования последовательностей информационных элементов (блоков) для многомерных кодовых комбинаций ЧЛБ (n_1, k_1; n_2, k_2)-кода с $k_1 = 4$ и $k_2 = |M| – (k_1 – 1)$, соответствующих моментам времени t_1, …, t_5

Используя изложенные подходы, можно получить цепь блоков цифровых данных, устойчивую к модификации, в том числе любому нарушению целостности (изменению) данных, содержащихся в блоках, а также нарушению хронологического порядка следования блоков цифровых данных в цепи. Указанные свойства являются крайне востребованными в настоящее время в цифровом пространстве. Учитывая, что цепь блоков цифровых данных получила их посредством процедуры кодирования ее элементов ЧЛБ (n, k)-кодом, будем в дальнейшем называть такую цепь ЧЛБ-цепью.

Проверочные элементы сформированных кодовых комбинаций для последующего хранения и использования могут записываться (помещаться) как дополнительные служебные данные непосредственно в k-тый информационный блок или записываться как самостоятельные блоки данных в отдельную служебную цепь, синхронизированную с исходной (порождающем их) цепью информационных блоков цифровых данных (элементов). При этом такая самостоятельная служебная цепь проверочных блоков (элементов) может храниться отдельно от основной информационной цепи. Кроме этого, в зависимости от решаемой практической задачи, возможен вариант, когда последовательность проверочных элементов (чисел) может оформляться в собственный служебный блок, который может добавляться как равноправный отдельный блок к единой последовательности информационных и проверочных блоков, в которой информационные и проверочные блоки будут чередоваться.

3 Факторы, оказывающие влияние на стойкость ЧЛБ-цепи к модификации

Стойкость к модификации ЧЛБ-цепей основана на следующих основных факторах:

- корректирующей способности выбранных ЧЛБ (n, k)–кодов;
- одновременном использовании каждого блока цифровых данных, представленного как информационный элемент, во многих (смежных с ним) кодовых комбинациях;
- способе выбора информационных элементов для формирования текущей последовательности

информационных элементов;

- возможности повышения стойкости к модификации цепи блоков цифровых данных за счет применения многомерного (каскадно-поступательного) кодирования информационных элементов;
- вариативности способов представления информационных элементов (блоков) как чисел;
- вариативности способов организации и использования проверочных элементов.

Рассмотрим некоторые факторы более подробно.

Известно, что ЧЛБ (n, k)–коды могут исправлять комбинации ошибок кратности $v \le [(d-1)/2]$ (где […] – знак целой части числа). При этом, ЧЛБ (n, k)–коды, полученные путем сочетания обычных (одномерных) ЧЛБ-кодов и называемые произведением кодов, итеративными или многомерными ЧЛБ-кодами, имеют минимальное кодовое расстояние, равное произведению кодовых расстояний составляющих их кодов. Например, для двумерного кода $d = d_1\ d_2$, где d_1 и d_2 – минимальные кодовые расстояния для кодов, составляющих двумерный код. Однако указанные выражения справедливы не во всех случаях, что отмечается в [8] при рассмотрении свойств прямого произведения кодов. Упоминание многомерных ЧЛБ (n, k)–кодов связано с тем, что представленный на рис. 2 метод построения одномерного линейно упорядоченного связного множества блоков цифровых данных (цепи) можно рассматривать как частный случай многомерного каскадно-поступательного (слово каскад здесь используется в значении последовательно соединенных однотипных объектов или конструкций) числового помехоустойчивого кодирования последовательности блоков цифровых данных.

При использовании предложенных способов кодирования высокая стойкость к модификации цепи блоков цифровых данных, связанных ЧЛБ (n, k)–кодом, обусловлена тем, что каждый блок цифровых данных, представленный как информационный элемент, используется одновременно во многих кодовых комбинациях. Например, в случае кодирования по методу, показанному на рис. 2, каждый информационный элемент, подвергающийся модификации, является информационным элементом одновременно k кодовых комбинаций (за исключением первых $k-1$ и последних $k-1$ информационных элементов, представляющих соответствующие блоки цифровых данных в цепи). Причем информационные элементы в каждой из этих кодовых комбинаций в свою очередь также взаимосвязаны с другими информационными элементами. Таким образом, помимо текущей кодовой комбинации каждый ее информационный элемент еще входит в последовательности информационных элементов $k-1$ смежных кодовых комбинаций. Это каскадно-поступательным образом обеспечивает одновременную, прямую или опосредованную, связь всех информационных элементов и соответствующих им блоков цифровых данных, входящих в конкретный момент времени в связную цепь блоков цифровых данных. Таким образом, модификация одного информационного блока цифровых данных нарушает целостность всей цепи, что позволяет обнаружить такой факт. При этом, в случае применения к блокам цепи принципов многомерного кодирования, как это показано, например, на рис. 3, количество смежных блоков, связанных с подвергающимся атаке блоком цепи, значительно возрастает, что многократно усложняет задачу потенциальному злоумышленнику.

Способ выбора информационных элементов для формирования текущей последовательности информационных элементов кодовой комбинации ЧЛБ (n, k)–кода, способ представления числами информационных элементов (блоков) цепи и способ организации и использования проверочных элементов (блоков) цепи во многом зависят от практической задачи, решаемой разработчиком технической системы, и также существенно влияют на стойкость ЧЛБ-цепи к модификации. Детальное рассмотрение конкретных способов и их влияния на свойства ЧЛБ-цепи является предметом для дальнейших исследований.

Заключение

Рассмотренные подходы к преобразованию массивов информации в цепь блоков цифровых данных, защищенную от модификации ЧЛБ (n, k)-кодом, названную автором ЧЛБ-цепью, способствуют решению актуальной в настоящее время задачи по обеспечению безопасности соответствующей информации, обрабатываемой в различных современных технических системах. По сравнению с другими известными подходами и технологиями, к числу которых можно отнести широко известную технологию блокчейн, в случае использования ЧЛБ-цепи открываются дополнительные возможности, связанные со свойствами ЧЛБ (n, k)-кодов. В частности, проверочные элементы кодовой комбинации содержат в себе определенные сведения об информационных

элементах, по которым они были вычислены, что позволяет определить модифицированные (искаженные) элементы этой кодовой комбинации, включая сами проверочные элементы. Это важное свойство позволяет не только вычислить в пределах корректирующей способности ЧЛБ-кода модифицированный элемент кодовой комбинации, т.е. автоматически восстановить искаженный блок, а также определить место и направление атаки, что представляется ценным при анализе и осуществлении противодействия возможным деструктивным воздействиям и минимизации их последствий. Кроме этого, достоинством ЧЛБ-кодов является то, что они допускают простое сложение массивов цифровых данных с другими однородными (согласованными) массивами цифровых данных с сохранением установленных свойств связности. Это способность обусловлена свойством ЧЛБ-кодов, при котором сумма кодовых комбинаций также является комбинацией ЧЛБ-кода.

Литература

1. Bitcoin: A Peer-to-Peer Electronic Cash System / Satoshi Nakamoto. URL: https://bitcoin.org/bitcoin.pdf (дата обращения: 20.11.2021).
2. *Терентьев, Андрей И.* Концептуальная схема (парадигма) технологии связных данных. Безопасность информационных технологий, [S.I], v. 28, n. 3, p. 65-72, sep. 2021. ISSN 2074-7136. Доступно на https://bit.mephi.ru/index.php/bit/article/view/1363. DOI: https://dx.doi.org/10.26583/bit.2021.3.05.
3. *Терентьев А.И.* Виды порождающих процедур связного множества цифровых данных / А.И. Терентьев // Гражданская авиация на современном этапе развития науки, техники и общества [Текст]: сборник тезисов докладов / Московский государственный технический университет гражданской авиации; редколлегия: Б. П. Елисеев (главный редактор) [и др.]. – М.: ИД Академии Жуковского, 2021. – 600 с.
4. *Терентьев А. И.* Метод обеспечения связности массивов цифровых данных посредством ЧЛБ (n, k)-кода / А. И. Терентьев // Управление развитием крупномасштабных систем (MLSD'2021) : Труды Четырнадцатой международной конференции, Москва, 27–29 сентября 2021 года / Под общей редакцией С.Н. Васильева, А.Д. Цвиркуна. – Москва: Институт проблем управления им. В.А. Трапезникова РАН, 2021. – С. 1558-1564. – DOI 10.25728/6491.2021.42.40.001. – EDN YGXSUD.
5. *Terentyev A.I.* Chain of Digital Data Blocks Linked by a Numeric Linear Block Correction Code, as an Alternative to the Blockchain Technology. Published in: 2021 14th International Conference Management of large-scale system development (MLSD). URL: https://doi.org/10.1109/MLSD52249.2021.9600226. DOI: 10.1109/MLSD52249.2021.9600226.
6. *Терентьев А. И.* Альтернативный технологии блокчейн метод построения упорядоченного связного множества блоков цифровых данных на основе числового корректирующего кода / А. И. Терентьев // FISP-2021: Фундаментальные проблемы информационной безопасности в условиях цифровой трансформации : Сборник докладов III Всероссийской научной конференции (с приглашением зарубежных ученых), Ставрополь, 30 ноября 2021 года. – Ставрополь: Северо-Кавказский федеральный университет, 2021. – С. 94-99. – EDN WYFVUX.
7. *Хохлов Г.И.* Числовые линейные блоковые корректирующие коды / Г.И. Хохлов // Электронная техника. сер.10. Микроэлектронные устройства. 1991, вып.2.
8. *Терентьев А.И.* Элементы теории и практики числовых линейных блоковых корректирующих кодов. – М.: Альтекс, 2000. – 204 с.: ил.