

Э. В. Фауре, к.т.н., доцент

e-mail: faureemil@gmail.com

Черкасский государственный технологический университет
б-р Шевченко, 460, г. Черкассы, 18006, Украина

ФАКТОРИАЛЬНОЕ КОДИРОВАНИЕ С ВОССТАНОВЛЕНИЕМ ДАННЫХ

В работе предложен и подробно рассмотрен метод факториального кодирования с восстановлением данных по перестановке. Предложенный метод направлен на комплексное решение задач криптографической защиты и защиты данных от ошибок канала связи. Факториальный код с восстановлением данных по перестановке является несистематическим кодом, предусматривающим замену информационной последовательности на перестановку чисел, вычисленную по всем информационным битам. Определены основные свойства предложенного кода, для которого выполнена оценка достоверности передачи, крипто- и имитостойкости. Изучены зависимости оценок вероятностей необнаруженной ошибки, энергетического выигрыша и скорости кода от длины информационного вектора на входе кодера. Выполнен сравнительный анализ известных факториальных кодов, приведена их классификация, даны рекомендации по применению.

Ключевые слова: факториальный код, перестановка, контроль целостности информации, криптозащита, помехоустойчивое кодирование, достоверность передачи, стойкость.

Постановка проблемы. В настоящее время при разработке и сопровождении информационно-телекоммуникационных систем и сетей необходимым условием является обеспечение конфиденциальности и контроля целостности передаваемой информации (КЦИ). Отметим, что КЦИ предусматривает комплексную защиту от навязывания ложных данных и обнаружение ошибок, вносимых каналом связи в процессе передачи сообщения.

Методы кодирования, обеспечивающие комплексное решение задач криптозащиты, имитозащиты и защиты данных от ошибок канала связи, позволяют повысить эффективность средств обработки информации за счет уменьшения вводимой избыточности и используемых вычислительных ресурсов, а разработка таких методов является актуальным направлением исследований.

Анализ источников и публикаций. Полученные в [1–3] результаты показывают эффективность применения факториального кодирования для задач КЦИ. При этом под факториальным кодированием информации понимается кодирование, использующее факториальную систему счисления для формирования кодового слова [3].

Так, в работе [2] исследованы свойства полного факториального кода (ПФК) [3], который в качестве проверочной части кодового слова использует перестановку чисел, опреде-

ляющуюся информационной последовательностью и алгоритмом кодирования. Здесь и далее под перестановкой порядка M понимается упорядоченный набор символов $\{0; 1; 2; \dots; M-1\}$.

Работа [2] направлена на исследование комбинированного факториального кода (КФК). В качестве проверочной части кодового слова КФК использует контрольную сумму циклического избыточного кода (CRC), вычисленную по проверочной части кодового слова ПФК.

Следует отметить, что ПФК и КФК являются систематическими избыточными кодами, которые направлены на решение задач КЦИ и не обеспечивают криптографической защиты передаваемой информации.

Целью данной работы является разработка и анализ метода факториального кодирования информации, реализующего в себе функции обнаружения ошибок в канале связи и криптозащиты данных, а также оценка его характеристик в системах передачи данных с решающей обратной связью (РОС). При этом оценке подлежат:

- скорость кода;
- вероятность не обнаруженной кодом ошибки;
- вероятность взлома кода методом «грубой силы».

Решение задачи. В силу того, что разрабатываемый код должен решать задачу

криптографической защиты передаваемого сообщения, он не может быть систематическим и содержать в кодовом слове информацию в открытом виде. Поэтому открытые данные с помощью хранящегося в тайне ключа должны быть преобразованы в шифртекст, который, помимо криптографической защиты, должен дополнительно обеспечить функции помехоустойчивого кодирования.

В качестве кода, удовлетворяющего предъявленным требованиям, предлагается факториальный код с восстановлением данных по перестановке.

Определение 1. Факториальным кодом с восстановлением данных по перестановке (ФКВД) называется несистематический код, предусматривающий замену информационной последовательности из k бит на перестановку чисел порядка M ($M! \geq 2^k$), вычисленную по всем k информационным битам.

Концепция метода. Как и в [2, 3], используем широко принятый [4, с. 601; 5, с. 232; 6, с. 361] подход к рассмотрению векторов над полем F_2 в виде элементов алгебры многочленов с коэффициентами из F_2 .

ФКВД (FCDR – Factorial Code with Data Recovery by Permutation) предусматривает замену информационной последовательности из k бит (вектора $A(x)$) на перестановку $R_{FCDR}(x)$ порядка M . При кодировании символов перестановки равномерным двоичным кодом ее длина $r = r_{FCDR} = M \cdot (\text{entier}(\log_2 M) + 1)$ бит, функция $\text{entier}(a)$ определяет наибольшее целое, меньшее, чем a . Перестановка в двоичном виде передается по каналу связи получателю. На станции приема выполняется обратное преобразование, что приводит к восстановлению k бит исходного блока данных и обеспечивает скорость кода:

$$v_{FCDR} = k/r_{FCDR}. \quad (1)$$

График зависимости скорости ФКВД от размера блока данных k на входе кодера показан на рис. 1. При этом значение M выбирается таким образом, чтобы $(M-1)! < 2^k \leq M!$.

Очевидно, что максимальная скорость ФКВД достигается при $(M! - 2^k) \rightarrow 0$ и $(\log_2 M - (\text{entier}(\log_2 M) + 1)) \rightarrow 0$. Поэтому с этой целью параметры кода целесообразно

выбирать таким образом, чтобы $\log_2 M \in \mathbb{Y}$, а длина блока данных k на входе кодера удовлетворяла условию $2^k \leq M! < 2^{k+1}$.

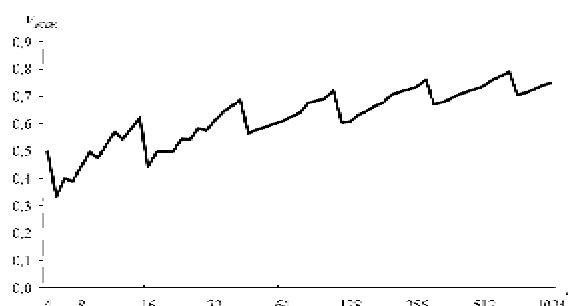


Рис. 1. График зависимости скорости ФКВД от размера блока данных на входе кодера

Заметим, что для обеспечения возможности восстановления данных по перестановке отображение $A(x) \leftrightarrow R_{FCDR}(x)$ должно быть биективным: каждому значению информационного вектора должна соответствовать одна перестановка, а каждой перестановке должно соответствовать одно значение информационного вектора. Следовательно, множества информационных векторов и перестановок должны быть равномошными. Поскольку $\log_2(M!) \notin \mathbb{Y}$ при $M \geq 3$, равенство $M! = 2^k$ справедливо только для $M = 2$ и $k = 1$. Поэтому в других случаях, которые представляют значительно больший интерес, необходимо соблюдать условие $M! > 2^k$, при этом $(M! - 2^k)$ перестановок должны быть запрещенными. В этом случае множество значений информационных векторов и множество разрешенных перестановок равномошны.

Соблюдение указанного условия может быть достигнуто следующим образом. Преобразование $f_{FCDR} : A(x) \rightarrow R_{FCDR}(x)$ заключается в преобразовании числа $A(x)$, представленного в двоичной системе счисления, в число A_F , представленное в факториальной системе счисления, с последующим его преобразованием в синдром S_F и перестановку π (см. [1, 7, 8, 9]) и кодированием ее символов равномерным двоичным кодом. Тогда $f_{FCDR} : A(x) \rightarrow A_F \rightarrow S_F \rightarrow \pi \rightarrow R_{FCDR}(x)$, а $f_{FCDR}^{-1} : R_{FCDR}(x) \rightarrow \pi \rightarrow S_F \rightarrow A_F \rightarrow A(x)$. Преобразования $A_F \leftrightarrow S_F \leftrightarrow \pi \leftrightarrow R_{FCDR}(x)$ для

определенной выше процедуры формирования перестановки по ее синдрому при фиксированной базовой перестановке $\pi(0)$ являются взаимно-однозначными. В свою очередь, преобразование $A_F \leftrightarrow A(x)$, предусматривающее преобразование числа A_F из факториальной в двоичную ($A(x)$) систему счисления и наоборот, также является взаимно-однозначным. Тогда численное значение $R_{FCDR}(x)$ не превышает значения $(2^k - 1)$, а все значения, большие $(2^k - 1)$, являются запрещенными.

Для разрушения статистической связи между информационной последовательностью $A(x)$ и соответствующей ей перестановкой $R_{FCDR}(x)$ последовательность $A(x)$ целесообразно подвергнуть скремблированию. Параметры скремблера могут храниться в секрете, что дополнительно повышает криптографическую стойкость преобразования.

Представление двоичного вектора $A(x)$ в факториальной системе счисления A_F решает задачу формирования проверочной части (перестановки) в зависимости от всех информационных символов, исключая при этом коллизии. Недостатком такого представления является невысокая скорость формирования перестановки за счет необходимости обработки двоичных чисел большой размерности. Альтернативой прямому преобразованию $A(x) \rightarrow A_F$ может служить следующая процедура:

- в память однократно записываются числа 2^j , $j \in [0, k-1]$, в факториальной системе счисления $\{A_F(2^{k-1}), A_F(2^{k-2}), \dots, A_F(2^1), A_F(2^0)\}$;

- по правилам сложения факториальных чисел $A_F = \sum_{i=0}^{k-1} a_i \cdot A_F(2^i)$, где a_i - i -й бит информационной последовательности. По полученному значению A_F легко определяется синдром S_F и перестановка π .

Такая процедура позволяет сократить время преобразования $A(x) \rightarrow A_F$, однако требует дополнительной памяти для хранения чисел $A_F(2^j)$, $j \in [0, k-1]$.

Оценка достоверности передачи. Все характеристики предложенного метода коди-

рования будем определять для простейшей системы передачи данных с РОС, где прямой канал - двоичный симметричный с переходной вероятностью p_0 ($q_0 = 1 - p_0$), обратный канал - идеальный, а символы, составляющие сообщение, являются элементами поля $F_2 = \{0; 1\}$.

Прежде всего, отметим, что в приемнике выполняется проверка корректности принятой из канала последовательности. Если в этой последовательности каждый из символов $\{0; 1; 2; \dots; M-1\}$ встречается ровно по одному разу, то полученная последовательность представляет собой перестановку и поступает на вход декодера. Если в принятой из канала последовательности пропущены или повторно применены какие-либо символы, то эта последовательность символов не является перестановкой и такой блок в системе с РОС подлжет переспросу.

Вероятность не обнаруженной декодером ФКВД ошибки $P_{ud}(FCDR, p_0)$ определяется вероятностью появления в канале связи такого вектора ошибок, который преобразует переданную перестановку в какую-либо из $(2^k - 1)$ других разрешенных перестановок. Если же все перестановки разрешены и обратное преобразование $f_{FCDR}^{-1}: R_{FCDR}(x) \rightarrow A(x)$ сюръективно, вероятность необнаруженной ошибки определяется вероятностью преобразования перестановки в какую-либо из $(M! - z)$ других перестановок, где z - количество перестановок, которые в результате обратного преобразования $f_{FCDR}^{-1}: R_{FCDR}(x) \rightarrow A(x)$ формируют переданную последовательность.

Поскольку $\begin{cases} 2^k - 1 \leq M! - 1, \\ M! - z \leq M! - 1, \end{cases}$ справедлива

оценка $P_{ud}(FCDR, p_0) \leq p_r$, где p_r вычислено в [2] (см. формулы (5-7), (10)) и определяет вероятность преобразования перестановки в другую перестановку при ее передаче по каналу связи:

$$P_{ud}(FCDR, p_0) \leq \sum_{i=1}^{m_1} f_{per}(2i) p_0^{2i} q_0^{r-2i} + \Delta_{per}(m_1), \quad (2)$$

где $f_{per}(0) = 1$; $f_{per}(2) \leq l_r \cdot M/2$;

$$f_{per}(4) \leq l_r \cdot M \cdot (l_r \cdot (M+8) - 10)/8, \quad (3)$$

$$\Delta_{per}(m_1) \leq e^{-\lambda} \cdot \frac{\lambda^{2(m_1+1)}}{(2(m_1+1))!} \times \frac{(2m_1+3)^2}{(2m_1+3)^2 - \lambda^2}, \quad (4)$$

$$l_r = \text{entier}(\log_2 M) + 1;$$

$$r = l_r \cdot M, \quad \lambda = r \cdot p_0;$$

m_1 выбирается таким образом, чтобы $m_1 > (\lambda - 3)/2$ и оценка $\Delta_{per}(m_1)$ не превышала заданной максимальной погрешности вычислений.

Пример. Оценим энергетический выигрыш ФКВД для некогерентного приема при $k = 716$, $p_0 = 10^{-3}$. Для $M = 128$ скорость кода $v_{FCDR} = 716/896 = 0.799$, а $P_{ud}(FCDR, p_0) \leq 4.814 \cdot 10^{-4}$. Энергетический выигрыш $\Delta P \geq 3.08$ дБ.

На рис. 2, а приведен график зависимости оценки вероятности необнаруженной ошибки от размера блока данных k на входе кодера в результате применения ФКВД при $p_0 = 10^{-3}$ и $M : (M - 1)! < 2^k \leq M!$. На рис. 2, б дополнительно точками отображены оценки вероятности необнаруженной ошибки, достигаемые в результате применения ПФК при идентичных ФКВД длине информационной части блока k и скорости кода.

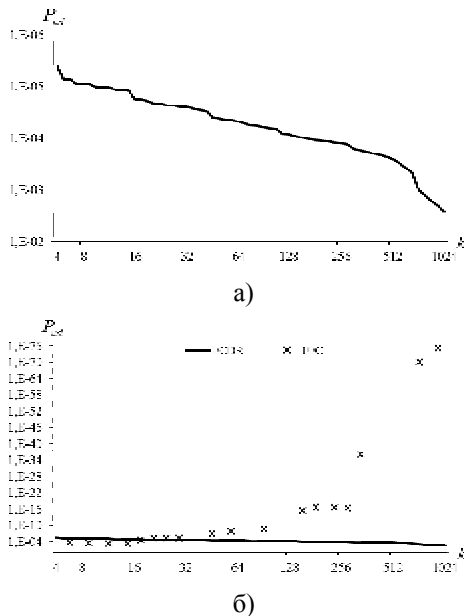


Рис. 2. Графики зависимостей оценок вероятностей необнаруженной ошибки от размера блока данных на входе кодера для ФКВД (а); ФКВД и ПФК (б) при $p_0 = 10^{-3}$

На рис. 3, а показан график зависимости оценки энергетического выигрыша ФКВД от размера блока данных k на входе кодера при $p_0 = 10^{-3}$ и $M : (M - 1)! < 2^k \leq M!$. На рис. 3, б дополнительно точками отображены оценки энергетического выигрыша ПФК при идентичных значениях длины информационной части блока k и скорости кода.

Рис. 2 и 3 свидетельствуют о том, что при малых значениях k ($k \leq 18$ для $p_0 = 10^{-3}$) энергетический выигрыш ФКВД превышает соответствующий энергетический выигрыш ПФК ($\Delta P_{FCDR} - \Delta P_{FFC} \leq 1.5$ дБ) при одинаковых скоростях кодов и кодировании символов перестановок равномерным двоичным кодом. В остальных случаях обнаруживающая способность ФКВД уступает обнаруживающей способности ПФК.

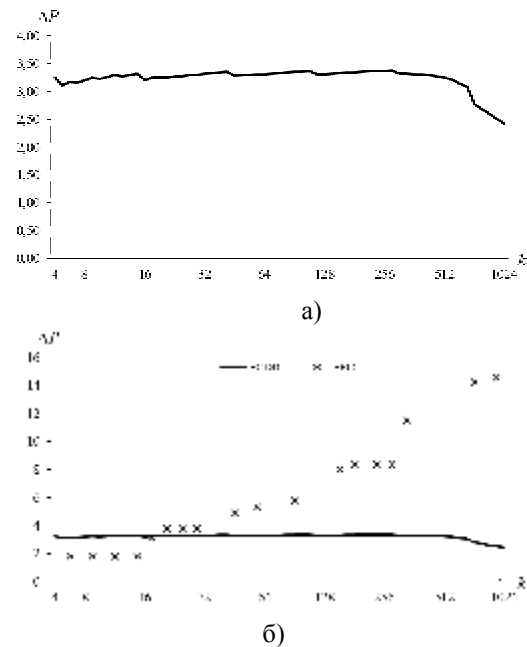


Рис. 3. Графики зависимостей оценок энергетического выигрыша от размера блока данных на входе кодера для ФКВД (а); ФКВД и ПФК (б) при $p_0 = 10^{-3}$

Оценка крипто- и имитостойкости ФКВД. Выполним количественную оценку стойкости ФКВД от несанкционированного чтения и/или навязывания ложных данных при атаке только на передаваемые данные и взломе методом «грубой силы» путем перебора множества значений ключевого пространства.

Вероятность взлома ФКВД методом грубой силы при однократной попытке подбора ключа $P_{UR}(FCDR) \leq (M!)^{-2}$.

Учитывая, что перестановка может быть легко модифицирована криптоаналитиком, ФКВД не обеспечивает имитозащиты сообщения.

Среднее время взлома системы КЦИ (системы защиты от НСД, подбора имитовставки) определяется выражением $T_{br} = 0.5 / (P_{br} \cdot N)$ с, где P_{br} – вероятность взлома системы КЦИ (системы защиты от

НСД, подбора имитовставки) для анализируемого кода, N – производительность компьютерной группировки, выполняющей процедуру взлома (ключей/с).

Классификация методов факториального кодирования. Выполним сравнительную оценку факториальных кодов и CRC-кода. Результаты анализа приведем в табл. 1.

Таблица 1

Свойства помехоустойчивых кодов

Код	Систематический	Помехоустойчивый	Криптоустойчивый	Имитостойкий	Самосинхронизирующийся
ПФК	+	+	-	+	+
КФК	+	+	-	+	-
ФКВД	-	+	+	-	+
CRC	+	+	-	-	-

Анализ свойств представленных помехоустойчивых кодов позволяет сформулировать следующие рекомендации по их применению:

1) при необходимости обеспечения КЦИ при ее передаче или хранении можно использовать ПФК или КФК. Причем, если не требуется самосинхронизации кода, более эффективно использовать КФК;

2) при необходимости совмещения свойств обнаружения ошибок в канале связи и защиты информации от несанкционированного доступа можно использовать ФКВД.

Выводы. Выполненная работа позволила разработать метод факториального кодирования с восстановлением данных по перестановке, который направлен на защиту информации от несанкционированного доступа и ошибок канала связи.

Свойства ФКВД:

– обеспечивает защиту от несанкционированного чтения, поскольку ключ формирования перестановки по информационной последовательности хранится в секрете;

– обеспечивает защиту от ошибок в канале связи;

– не обеспечивает имитозащиту;

– при одинаковых длине кодовой комбинации и скорости кода обнаруживающая способность ФКВД при малой длине блока данных на входе кодера ($k \leq 18$ для $p_0 = 10^{-3}$) выше, чем ПФК, однако ниже CRC-кода;

– обладает свойством самосинхронизации, которое обусловлено тем, что символы

$\{0; 1; \dots; M-1\}$ встречаются в перестановке ровно по одному разу, а их сумма равна $\sigma = 0,5M(M-1)$;

– исключает коллизии.

Список литературы

1. Фауре Э. В. Метод формирования имитовставки на основе перестановок [Электронный ресурс] / Э. В. Фауре, В. В. Швыдкий, В. А. Щерба // Захист інформації. – 2014. – № 4, т. 16. – С. 334–340. – Режим доступа: <http://jrn1.nau.edu.ua/index.php/ZI/article/view/334/8755>
2. Фауре Э. В. Контроль целостности информации на основе факториальной системы счисления / Э. В. Фауре, В. В. Швыдкий, В. А. Щерба // Journal of Qafqaz University. Mathematics and computer science. – 2016. – № 2, т. 4. – (В печати).
3. Фауре Э. В. Комбинированное факториальное кодирование и его свойства / Э. В. Фауре, В. В. Швыдкий, В. А. Щерба // Радіоелектроніка, інформатика, управління. – 2016. – № 3. – (В печати).
4. Лидл Р. Конечные поля : в 2 т. / Р. Лидл, Г. Нидеррайтер ; пер. с англ. под ред. Нечаева В. И.]. – М. : Мир, 1988. – Т. 2. – 822 с. – (Редакция литературы по математическим наукам).
5. Питерсон У. Коды, исправляющие ошибки / У. Питерсон, Э. Уэлдон ; пер. с англ. под ред. Р. Л. Добрушина, С. И. Самойленко] – М. : Мир, 1976. – 590 с. – (Редакция литературы по новой технике).

6. Прокис Д. Цифровая связь / Джон Прокис ; пер. с англ. под ред. Д. Д. Кловского]. – М. : Радио и связь, 2000. – 800 с.
7. Фауре Э. В. Метод формирования воспроизводимой непредсказуемой последовательности перестановок [Электронный ресурс] / Э. В. Фауре, В. В. Швидкий, А. И. Щерба // Безпека інформації. – 2014. – № 3, т. 20. – С. 253–258. – Режим доступа : <http://jrn1.nau.edu.ua/index.php/Infosecurity/article/view/7552/8608>
8. Пат. 106668 Україна, МПК G06F 7/58 (2006.01). Спосіб формування випадкової послідовності перестановок / Фауре Е. В., Швидкий В. В., Щерба А. І. ; заявник та патентовласник ЧДТУ. – № а201505933; заявл. 16.06.2015; опубл. 10.05.2016, Бюл. № 9.
9. Пат. 106669 Україна, МПК G06F 21/64 (2013.01). Спосіб формування імітовставки / Фауре Е. В., Швидкий В. В., Щерба А. І. ; заявник та патентовласник ЧДТУ. – № а201505934; заявл. 16.06.2015; опубл. 10.05.2016, Бюл. № 9.
3. Faure, E. V., Shvydkyi, V. V. and Shcherba, V. A. (2016). Combined factorial coding and its properties. *Radioelektronika, informatyka, upravlinnya*, (3), (in print) [in Russian].
4. Lidl, Rudolf, and Niederreiter, Harald (1997) Finite fields. Cambridge: Cambridge UP.
5. Peterson, W. W. and Weldon, E. J. (1972) Error-correcting codes. 2nd ed. Cambridge (Mass.); London : M.I.T., 1972.
6. Proakis, John G. (2001) Digital communications. Boston: McGraw-Hill.
7. Faure, E. V., Shvydkyi, V. V. and Shcherba, A. I. (2014). Method of forming reproducible and unpredictable sequence of permutations. *Bezpeka Informatsiyi*, 3 (20), pp. 253–258, available at: <http://jrn1.nau.edu.ua/index.php/Infosecurity/article/view/7552/8608>
8. Faure, E. V., Shvydkyi, V. V. and Shcherba, A. I. (2016) Patent of Ukraine 106668. MPK G06F 7/58 (2006.01) The method of formation of random permutations sequence. The owner Cherkasy State Technological University, No. a201505933, stated 16.06.2015; publ. 10.05.2016, Bul. No. 9.
9. Faure, E. V., Shvydkyi, V. V. and Shcherba, A. I. (2016) Patent of Ukraine 106669. MPK G06F 21/64 (2013.01). The method of message authentication code formation. The owner Cherkasy State Technological University, No. a201505934, stated 16.06.2015; publ. 10.05.2016, Bul. No. 9.

References

1. Faure, E. V., Shvydkyi, V. V. and Shcherba, V. A. (2014). Method of message authentication code formation based on permutations. *Zakhyst informatsiyi*, (4), pp. 334–340, available at: <http://jrn1.nau.edu.ua/index.php/ZI/article/view/334/8755>
2. Faure, E. V., Shvydkyi, V. V. and Shcherba, A. I. (2016). Information integrity control

E. V. Faure, *Ph.D., associate professor*,
Cherkasy State Technological University
Shevchenko blvd, 460, Cherkasy, 18006, Ukraine
e-mail: faureemil@gmail.com

FACTORIAL CODING WITH DATA RECOVERY

Introduction. *Methods of coding that provide a comprehensive solution of cryptographic protection, protection against intentional alteration of data, and data protection against communication channel errors can improve the efficiency of information processing tools by reducing included redundancy and used computing resources. The development of such methods is a topical area of current research.*

The purpose of this study *is to develop and analyze the method of factorial coding with data recovery that realizes functions of communication channel error detection and cryptographic protection of information, and to assess its properties in data transmission systems with decision feedback.*

The main material. Factorial code with data recovery by permutation is a non-systematic code that provides a replacement of information sequence into permutation of the numbers computed by all information bits. In the article, the main properties of the proposed code are defined. The transmission reliability, cryptographic strength and strength against intentional alteration of data are evaluated. Dependences of undetected error probability assessments, energy gain and code rate from information vector length at the encoder input are studied. A comparative analysis of known factorial codes is done. Recommendations for the use of factorial coding are given.

Conclusions. The performed work extends the scientific and technical base of factorial methods and means of data coding. The proposed factorial code with data recovery: provides protection against unauthorized reading, provides protection against channel errors, does not provide protection against intentional alteration of data, has higher detection ability for small data block length on encoder input as compared with a full factorial code, has the property of self-synchronization, and excludes collisions.

Keywords: factorial code, permutation, information integrity control, cryptographic protection, error control coding, transmission accuracy, strength.

Рецензенти: В. М. Рудницький, д.т.н., професор,
С. В. Голуб, д.т.н., професор