УДК 519.689

А.А. Смагин, П.И. Смикун

МОДЕЛИРОВАНИЕ ПРОЦЕССА ПЕРЕДАЧИ ДАННЫХ НА ОСНОВЕ СИНХРОНИЗАЦИИ ПЕРЕДАЮЩЕГО И ПРИЕМНОГО УСТРОЙСТВ

***Смагин Алексей Аркадьевич**, доктор технических наук, профессор, окончил радиотехнический факультет Ульяновского политехнического института. Заведующий кафедрой телекоммуникационных технологий и сетей Ульяновского государственного университета. Имеет свыше 100 статей, изобретения, монографию в области разработки информационных систем различного назначения. [Тел.: (8422) 32-01-00 / E-mail: AlSmagin@ulsu.ru].*

***Смикун Петр Иванович**, окончил факультет автоматики и вычислительной техники Таганрогского радиотехнического института. Начальник научно-исследовательского отделения - заместитель главного конструктора ФНПЦ ОАО «НПО «Марс». Имеет статьи в области разработки информационных систем различного назначения. [Тел.: (8422) 26-28-98 / E-mail: smikun@mail.ru].*

Аннотация

Изложен принцип трансляции данных по каналам связи, реализующий идею сжатия и одновременного достижения помехоустойчивости. Определена вероятность условий, при которых достижим этот эффект.

Ключевые слова: матрично-ранговое кодирование, помехоустойчивость, сжатие данных.

Abstract

The article gives an account of a principle of data translation via communications channels, which implements an idea of compression and simultaneous jam-immunity achievement. It defines a probability of conditions when the effect is achievable.

Key words: matrix and rank encoding, jam-immunity, data compression.

Как известно, функционирование системы «передатчик-приемник» имеет множество аспектов, к числу которых относятся скорость передачи информации и помехоустойчивость. При этом, как правило, помехоустойчивость достигается за счет введения дополнительной избыточности, что приводит к уменьшению скорости передачи кода. А это, в свою очередь, замедляет процесс передачи информации за счет загруженности канала связи избыточными данными. Существующий уровень технического обеспечения канала связи позволяет реализовать идею сжатия и одновременного достижения помехоустойчивости, что приводит к существенному увеличению скорости передачи данных.

В настоящем материале излагается принцип трансляции данных по каналам связи, дающий подобный эффект. Достижение этого эффекта возможно при определенных, имеющих высокую вероятность, условиях, характеризующих источник информации, и зависит от надежности синхронизации приемного и передающего устройств.

Пусть имеется некоторая бинарная последовательность, подлежащая передаче по каналам связи. Пусть p — вероятность появления единицы в данной последовательности. Предлагаемый подход для передачи информации с использованием метода матрично-рангового кодирования заключается в следующем: разбиваем бинарную последовательность на n блоков длины L. Предположим, что блоки длины L не повторяются, т.е. не существует двух или более одинаковых блоков. В утверждении приводится вероятность подобной ситуации.

Утверждение. *Вероятность $\widehat{p}$ того, что среди двоичных блоков длины L не существует повторяющихся, равна*

$$\widehat{p} = n!\sum_{k=1}^{C^{n}_{2^{L-1}}} \prod_{i=1}^{2^L} p_i^{x_i^k} q_i^{1-x_i^k}, \tag{1}$$

где x_i^k - i-я компонента равномерного кода длины 2^L числа k, полученного методом матрично-рангового кодирования с рангом n;

$p_i = p^{Ri}q^{L-Ri}$, R^i - количество единиц в двоичном представлении числа i, вычисляемое по рекурсивной формуле $R_i = g(x)$, где функция $g(x)$ определяется следующим образом:

$$\begin{cases} g(0)=0 \\ g(1)=1 \\ g(x)=I\left(\left[\dfrac{x}{2^{[\log_2 x]}}\right]\right)+g\left(x-2^{[\log_2 x]}\right). \end{cases}$$

Доказательство. Пронумеруем все возможные двоичные последовательности длины L при помощи перевода в десятичный код с использованием стандартной процедуры. Пусть i -десятичное число, полученное в результате операции декодирования, R_i - ранг i-го блока длины L. Определим рекурсивную функцию $g(x)$ следующим образом:

$$g(0)=0, \quad g(1)=1,$$

$$g(x)=I\left(\left[\frac{x}{2^{[\log_2 x]}}\right]\right)+g\left(x-2^{[\log_2 x]}\right),$$

тогда ранг i-го блока длины L может быть вычислен при помощи построенной функции:

$$R_i = g(i).$$

Очевидно, исходя из независимости появления символов в блоке, согласно теореме умножения вероятностей, вероятность появления блока длиной L, содержащего R_i единиц, будет равна

$$p_i = p^{R_i}(1-p)^{L-R_i}.$$

Таким образом, имеется 2^L исходов с соответствующими вероятностями $p_i, i = 1, ..., 2^L$. Пусть $x=\left(x_1^k, x_2^k, \ldots, x_{2^L}^k\right)$ - произвольный двоичный вектор, содержащий n единиц (о значении индекса k будет сказано в последующих рассуждениях). Воспользовавшись полиномиальной формулой в схеме Бернулли, получим вероятность конкретного набора $x=\left(x_1^k, x_2^k, \ldots, x_{2^L}^k\right)$ равную $n!\sum_{k=0}^{C^n_{2^{L-1}}}\prod_{i=1}^{2^L} p_i^{x_i^k} q_i^{1-x_i^k}$. Далее, для того, чтобы пронумеровать все векторы x, содержащие n единиц, воспользуемся методом матрично-рангового кодирования. Согласно свойству монотонного неубывания длины кода по кодируемому десятичному числу при фиксированном ранге, все возможные комбинации бинарных блоков длины L, содержащие n единиц, будут являться кодами чисел от 0 до $C^n_{2^{L-1}}$ с рангом n, полученными методом матрично-рангового кодирования. Следовательно, искомая вероятность может быть найдена по формуле

$$\hat{p} = n!\sum_{k=0}^{C^n_{2^{L-1}}}\prod_{i=1}^{2^L} p_i^{x_i^k} q_i^{1-x_i^k},$$

где x_i^k - i-я компонента равномерного кода длины 2^L числа k, полученного методом матрично-рангового кодирования с рангом n. Утверждение доказано.

Предположим, что, согласно закону больших чисел, при росте длины L в каждом блоке от $[Lp - \Delta]$ до $[Lp + \Delta]$ единиц ранг варьируется в диапазоне ($[Lp - \Delta]$, $[Lp + \Delta]$). Сразу заметим, что $[Lp - \Delta] \geq 1$ и $[Lp + \Delta] \leq L.$ Объединив эти два неравенства, получим

$$\Delta \leq min\,(Lp - 1, Lq). \quad (2)$$

Всего кодов длины L, полученных методом матрично-рангового кодирования, 2^{L-1} (нулевой байт не входит в допустимое множество, и, кроме того, код меньшей длины может быть корректно приведен к равномерному коду длины L). Это следует, например, из свойства монотонности длины кода по кодируемому десятичному числу при фиксированном ранге. При ограничении на диапазон ранга количество кодов сокращается до $f_L(\Delta)$.

$$f_L(\Delta)=\sum_{i=[L_p-\Delta]}^{[L_p+\Delta]} C_{L-1}^{i-1}. \quad (3)$$

Очевидно, что функция $f_L(\Delta)$ является монотонно возрастающей.

Итак, существует $f_L(\Delta)$ потенциально возможных кодов. Пронумеруем данные коды произвольным образом и зафиксируем данный порядок как первоначальный. Каждый блок, полученный в результате разбиения исходной последовательности, является кодом некоторого числа N с рангом R. Следовательно, согласно данной нумерации ему может быть присвоен некоторый номер (см. таблицу 1).

Таблица 1

Нумерация блоков исходной последовательности

Номер блока	1	2	$\ldots$	n
Номер согласно нумерации	x_1	x_2	$\ldots$	x_n

Осуществим перестановку множества $\{1, 2, ..., f_L(\Delta)\}$ так, чтобы

$$x_1 \to 1; \;\; x_2 \to 2; \;\; \ldots \;\; x_n \to n.$$

Сразу отметим, что для корректной передачи количество не должно превышать $f_L(\Delta)!$, то есть $n \leq f_L(\Delta)!$. Пусть M - код необходимой перестановки по алгоритму генерации перестановок. Очевидно, что

$$0 \leq M \leq f_L(\Delta)!. \quad (4)$$

Идея передачи информации заключается в синхронизации приемного и передающего устройств. Пусть i - показания таймера передатчика. В мантиссе времени отводится z разрядов для заполнения их информацией о числе M. Число M требует, очевидно, $[lgM]$ десятичных разрядов.

Далее, начиная со времени t, осуществляется n посылок с произвольным интервалом (или только одна посылка, а число блоков передается другим способом или заранее согласовано). Декодирование осуществляется по времени t первой посылки и количеству n блоков (длина L заранее задается приемным и передающим устройствами). Для этого по t восстанавливается M (код перестановки), затем непосредственно сама перестановка. Определяются прообразы $1, 2, \ldots, n$. Это числа $x_1, x_2, \ldots, x_n$. Затем по первоначальной нумерации восстанавливаются сами блоки.

Пусть $\tilde{p}$ - надежность передачи, характеризующая четкость синхронизации установленных z разрядов таймеров. Зададимся вопросом: сколько необходимо импульсов в установленное время, чтобы с вероятностью, большей или равной p`, по принципу «большинства голосов» была осуществлена правильная передача. Обозначим через k необходимое количество импульсов. Тогда вероятность $P(\tilde{p})$ правильной передачи по принципу «большинства голосов» в предположении о выполнении условий схемы Бернулли может быть определена по формуле:

$$P(\tilde{p})=\sum_{i=\left|\frac{k}{2}\right|+1}^{k} C_k^i \tilde{p}^i (1-\tilde{p})^{k-i}. \qquad (5)$$

Оценим возможность выполнения условия $P(\tilde{p}) \geq p`$. В таблице 2 приведено соответствие различных значений переменной количеству импульсов.

Очевидно, что нужно выбирать наименьшее k, удовлетворяющее условию $P(\tilde{p}) \geq p`$ при заданных $\tilde{p}$ и $p`$.

Очевидным образом следует, что при больших значениях переменной количества импульсов

Таблица 2

Соответствие значений переменной k количеству импульсов

k	$P(\tilde{p})$
3	$C_3^2 \tilde{p}^2 (1-\tilde{p})+C_3^3 \tilde{p}^3$
4	$C_4^3 \tilde{p}^3 (1-\tilde{p})+C_4^4 \tilde{p}^4$
5	$C_5^3 \tilde{p}^3 (1-\tilde{p})^2+C_5^4 \tilde{p}^4 (1-\tilde{p})+C_5^5 \tilde{p}^5$
6	$C_6^4 \tilde{p}^4 (1-\tilde{p})^2+C_6^5 \tilde{p}^5 (1-\tilde{p})+C_6^6 \tilde{p}^6$
7	$C_7^4 \tilde{p}^4 (1-\tilde{p})^3+C_7^5 \tilde{p}^5 (1-\tilde{p})^2+ +C_7^6 \tilde{p}^6 (1-\tilde{p})+C_7^7 \tilde{p}^7$

для случая $\tilde{p} > ½$ вероятность $P(\tilde{p})$ с большой скоростью асимптотически стремится к единице. Тем самым это гарантирует выполнение условия $P(\tilde{p}) \geq p`$ при $\tilde{p} > ½$. Рассмотрим более подробно случай неограниченного увеличения k. Воспользовавшись теоремой Лапласа, получим:

$$P(\tilde{p})=1-\varphi\left(\frac{k/2-k\tilde{p}}{\sqrt{k\tilde{p}(1-\tilde{p})}}\right), \qquad (6)$$

где

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{t^2}{2}}dt. \qquad (7)$$

Далее для выполнения условия $P(\tilde{p}) \geq p`$ решим неравенство:

$$1-\varphi\left(\sqrt{k}\,\frac{1/2-\tilde{p}}{\sqrt{\tilde{p}(1-\tilde{p})}}\right)\geq p`. \qquad (8)$$

Из свойства монотонного возрастания функции распределения для стандартного нормального распределения правомерно следующее действие:

$$\sqrt{k}\,\frac{1/2-\tilde{p}}{\sqrt{\tilde{p}(1-\tilde{p})}}\leq \varphi^{-1}(1-p`).$$

Следовательно, при $p` > ½$ и $\tilde{p} > ½$ верно

$$k\leq\frac{4\tilde{p}(1-\tilde{p})}{(1-2\tilde{p})^2}u_{1-p`}^2, \qquad (9)$$

где $u_{1-p`}^2$ - квантиль порядка $1 - p`$ стандартного нормального распределения.

Заметим, что при увеличении Δ эффект от использования метода матрично-рангового кодирования уменьшается из-за увеличения $f_L(\Delta)$, а следовательно, возможного значения M, что требует большего количества разрядов в мантиссе времени для передачи информации. А разряды более высокого порядка, как известно, дают большую погрешность при синхронизации. Более того, при увеличении Δ до предельного значения разработанная процедура в виде частного случая становится процедурой использования стандартного алгоритма нумерации бинарных последовательностей путем представления их десятичным кодом.

Количество разрядов мантиссы для синхронизации работы таймеров составит не более

$$\lceil lg\, f_L(\Delta) \rceil = \left\lceil \sum_{i=1}^{f_L(\Delta)} lg\, i \right\rceil. \quad (10)$$

Таким образом, при уменьшении Δ увеличивается эффект от использования метода матрично-рангового кодирования в $\frac{2^L}{f_L(\Delta)}$ раз (здесь в качестве сравнительной эффективности рассматривается отношение количества потенциально возможных кодов).

СПИСОК ЛИТЕРАТУРЫ

1. Смагин А.А., Терентьева Ю.Ю. Способ преобразования дискретной информации. Фундаментальные проблемы математики и механики // Ученые записки УлГУ. - Ульяновск, УлГУ. - 1996. - Вып. 1. - Часть 2. - С. 101-108.

2. Смагин А.А., Смикун П.И., Терентьева Ю.Ю. Алгоритм помехоустойчивой передачи потока данных // Известия научного центра РАН. — 2008. - Т.4. - С. 96-98.