

МРНТИ 89.25.21
УДК 629.7.05.001

<https://doi.org/10.51889/2021-3.1728-7901.12>

*Б.-Б.С. Есмагамбетов**

Южно-Казахстанский университет им. М. Ауэзова, Шымкент, Казахстан

**email: bulatbatyr@mail.ru*

ПОМЕХОУСТОЙЧИВОСТЬ В ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ СО СЖАТИЕМ ДАННЫХ

Аннотация

При обработке в информационных системах огромных потоков данных отдельные их измерения или целые группы измерений могут быть искажены или утеряны вследствие разных причин. Восстановление сжатых данных при передаче по каналам связи сопровождается погрешностями, связанными с искажением информационной и служебной частей сообщений из-за наличия помех в канале передачи. К этим погрешностям добавляются ошибки, обусловленные квантованием передаваемых реализаций по уровню и дискретизацией по времени. Исследования методов повышения помехоустойчивости как при передаче, так и при восстановлении измеряемых данных является актуальной задачей при проектировании информационно-измерительных систем.

В статье рассмотрены непараметрические методы оценивания вероятностных характеристик случайных процессов. Отличительной особенностью непараметрических методов является ранжирование данных, измеряемых на интервале наблюдения. Показано, что ранжирование данных на передающей стороне информационно-измерительной системы дает возможность коррекции ошибок и сбоев основанной на строгой монотонности ранжированного ряда кодов. Также проведено исследование погрешности восстановления непрерывных реализаций с учетом искажений сжатых данных в канале связи. Полученные результаты свидетельствуют о том, что применение сложных алгоритмов сжатия нецелесообразно, так как разница в погрешности восстановления нестационарных сообщений между простейшим алгоритмом и достаточно сложным становится пренебрежимо малой. В статье приведены результаты оценки погрешностей восстановления для различных методов сжатия данных.

Ключевые слова: сжатие данных, помехоустойчивость, случайный процесс, существенные отсчеты, избыточность, непараметрические статистики.

Аңдатпа

Б.С. Есмагамбетов

М. Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан

ДЕРЕКТЕРДІ СЫҒУ АҚПАРАТТЫҚ-ӨЛШЕУ ЖҮЙЕЛЕРІНДЕГІ БӨГЕУЛ ОРНЫҚТЫЛЫҚ

Ақпараттық жүйелерде деректердің орасан зор ағындарын өңдеу кезінде олардың жекелеген өлшемдері немесе тұтас өлшеу топтары әртүрлі себептердің салдарынан бұрмалануы немесе жоғалуы мүмкін. Байланыс арналары арқылы беру кезінде сығылған деректерді қалпына келтіру хабарлардың ақпараттық және қызметтік бөліктерін бұрмалаумен байланысты қателіктермен қоса беріледі. Бұл қателіктерге берілетін сатуларды деңгей бойынша кванттаумен және уақыт бойынша дискреттеумен негізделген қателер қосылады. Өлшенетін деректерді беру кезінде де, қалпына келтіру кезінде де кедергіге төзімділікті арттыру әдістерін зерттеу ақпараттық-өлшеу жүйелерін жобалау кезіндегі өзекті міндет болып табылады.

Мақалада кездейсоқ процестердің ықтимал сипаттамаларын бағалаудың параметрлік емес әдістері қаралды. Параметрлік емес әдістердің ерекшелігі бақылау аралығында өлшенетін деректерді саралау болып табылады. Ақпараттық-өлшеу жүйесінің таратушы жағындағы деректерді саралау сараланған бірқатар кодтардың қатаң монотондылығына негізделген қателер мен іркілістерді түзетуге мүмкіндік беретіні көрсетілген. Сонымен қатар, байланыс арнасындағы сығылған деректердің бұрмалануын ескере отырып, үздіксіз іске асыруды қалпына келтіру қателіктеріне зерттеу жүргізілді. Алынған нәтижелер сығымдаудың күрделі алгоритмдерін қолдану орынсыз екенін көрсетеді, себебі қарапайым алгоритм мен айтарлықтай күрделі арасындағы стационарлық емес хабарламаларды қалпына келтіру қателіктеріндегі айырмашылық елеусіз болады. Мақалада деректерді қысудың әртүрлі әдістері үшін қалпына келтіру қателіктерін бағалау нәтижелері келтірілген.

Түйін сөздер: деректерді сығу, бөгеу орнықтылық, кездейсоқ процесс, елеулі есептеулер, артықтық параметрлік емес статистикалар.

Abstract

NOISE IMMUNITY IN DATA COMPRESSION INFORMATION AND MEASUREMENT SYSTEMS

Yesmagambetov B.-B.S.

South Kazakhstan University named after M. Auezov, Shymkent, Kazakhstan

When processing huge data streams in information systems, individual measurements or whole groups of measurements can be distorted or lost due to various reasons. Recovery of compressed data during transmission on communication channels is accompanied by errors related to distortion of information and service parts of messages due to presence of interference in transmission channel. To these errors are added errors caused by quantization of the transmitted implementations by level and time sampling. Research on methods of increasing noise immunity both during transmission and during recovery of measured data is an urgent task in the design of information and measurement systems.

The article considers non-parametric methods of estimating probabilistic characteristics of random processes. A distinctive feature of non-parametric methods is the ranking of data measured at the observation interval. It is shown that ranking of data on transmitting side of information-measuring system enables correction of errors and failures based on strict monotony of ranked number of codes. Also, the error of recovery of continuous implementations taking into account distortions of compressed data in the communication channel was investigated. The obtained results indicate that the use of complex compression algorithms is impractical, since the difference in the error in the restoration of non-stationary messages between the simplest algorithm and the rather difficult one becomes negligible. The article presents the results of estimating recovery errors for various data compression methods.

Keywords: data compression, noise immunity, random process, significant counts, redundancy, non-parametric statistics.

Введение

В информационно-измерительных системах исследования, связанные с оценкой действия помех, являются чрезвычайно актуальными ввиду влияния последних на качество передаваемой и принимаемой информации. В системах со сжатием данных, особенно в бортовых телеметрических системах летательных аппаратов, актуальность вопросов помехоустойчивости передаваемых сообщений еще более усиливается в связи с наличием погрешностей, накладываемых самими методами сжатия.

Апертурные методы сжатия данных не эффективны при наличии помех, так как практически все сбои фиксируются в качестве существенных отсчетов, что приводит к существенным искажениям результатов при их восстановлении. По сравнению с ними необратимые методы сжатия с использованием непараметрических методов теории решений дают существенно лучшие результаты как по коэффициенту сжатия, так и по точностным характеристикам восстановления и трудоемкости вычислительных операций. Главное их достоинство заключается в устойчивости методов к различного рода шумам.

Методология исследования

Простейшая модель нестационарного процесса имеет вид

$$y(t) = x(t) + h(t),$$

где $y(t)$ - измеряемый процесс, $x(t)$ - случайная стационарная составляющая, $h(t)$ - нестационарный тренд (сигнал) [1].

Процедура обработки заключается в разделении исходной реализации $y(t)$ на две составляющие в соответствии с выше приведенным выражением. Основные этапы алгоритма следующие. Исходная реализация делится на последовательность отрезков $\tau_{ст}$, на которых она стационарна по математическому ожиданию. Степень стационарности на отрезке регулируется априори задаваемым уровнем значимости α , величина которого показывает, что на данном отрезке математическое ожидание меняется не более, чем на ϵ - отклонение. Выход его за границы ϵ -отклонения соответствует концу текущего отрезка стационарности и началу нового отрезка стационарности $\tau_{ст}$. Оценка среднего значения процесса на каждом отрезке стационарности получается путем присвоения ему значения процесса в середине отрезка стационарности [2].

Таким образом, получаем оценку тренда $h^*(t)$. Случайный стационарный остаток $x^*(t)$ может быть получен как $x^*(t) = y(t) - h^*(t)$.

Для исследования процедуры использовались нестационарные модели вида

$$y(t) = \exp(i/100) + x(t),$$

$$y(t) = 5 \sin(0.05i) + x(t),$$

где $x(t)$ – стационарная составляющая, имеющая нормальное распределение. Составляющие, взятые в качестве тренда в этих моделях, описываются монотонными и колебательными функциями. На рис.1 приведены результаты операции выделения нестационарной составляющей для монотонного тренда.

Пользуясь терминами сжатия данных, за существенные точки можно принять номера отсчетов, соответствующие серединам отрезков стационарности (иногда за существенные точки принимают номера отсчетов, соответствующие границам отрезков – правым или левым).

Таким образом, сокращение избыточности происходит за счет передачи существенных точек исходного процесса $y(t)$, характеризующих тренд. В случае, если необходимо иметь информацию о стационарном остатке, то передаются его статистические характеристики.

Для этого достаточно проранжировать полученные на интервале квазистационарности данные в порядке их возрастания или убывания. Для оценивания среднеквадратического отклонения можно использовать крайние порядковые статистики [3]:

$$\sigma = \gamma(X^{(N)} - X^{(1)})$$

или, с учетом сбойных наблюдений:

$$\sigma = \gamma(X^{(N-1)} - X^{(2)}),$$

где γ – коэффициент, зависящий от объема наблюдений на интервале квазистационарности. Рекомендуется выбирать $\gamma = 1/3$ при $n < 12$ и $\gamma = 1/4$ при $n > 12$ [4].

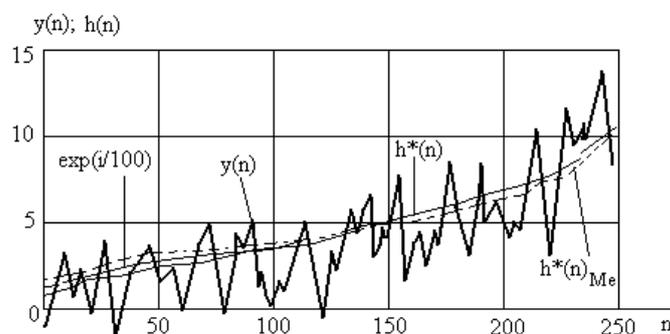


Рисунок 1. Выделение нестационарной составляющей

Для оценивания функции распределения и корреляционной функции можно применить формулы использующие ранговые статистики [5,6]:

$$F(x) = R/(N + 1),$$

$$\rho_k(j) = K(N) \sum R_i - 1.$$

Здесь R – ранг элемента, K – постоянная Кендалла, N – объем наблюдений на интервале квазистационарности.

Существенный выигрыш в вычислительных затратах получается, если на одном интервале стационарности оценивать несколько вероятностных характеристик, так как при их вычислении используется один и тот же ряд ранжированных наблюдений.

Ранжированные на передающей стороне данные дают возможность коррекции ошибок и сбоев на приемной стороне информационно-измерительной системы основанной на строгой монотонности ранжированного в (пределах пакета) ряда $x_1, x_2, \dots, x_R, \dots, x_N$ кодов.

Ее можно пояснить с помощью примера, приведенного на рис.2. Здесь через x^*_R обозначен код (ранга R) с ошибкой передачи, а через \bar{x}_R – его исправленное значение (рис. 2).

При этом очевидно, что максимальное значение погрешности передачи с применением коррекции не превосходит в каждом одиночном случае величины $\Delta_R = X_R - X_{R-1}$, а среднеквадратическая погрешность – величины

$$\varepsilon_R = \sqrt{0,5} \sum \Delta_R^2.$$

Отметим, что передача порядковых статистик, несущих информацию о различных характеристиках процесса, при их расположении в пакете в порядке возрастания или убывания рангов представляет частный случай предложенного способа.

Другой важной характеристикой работы алгоритмов является их помехоустойчивость не только к наличию шумовой составляющей, но и к сбойным точкам. Этот вопрос исследовался экспериментально путем добавления к вышеприведенным моделям сбойных точек, число и значения которых менялись случайным образом по соответствующей программе.

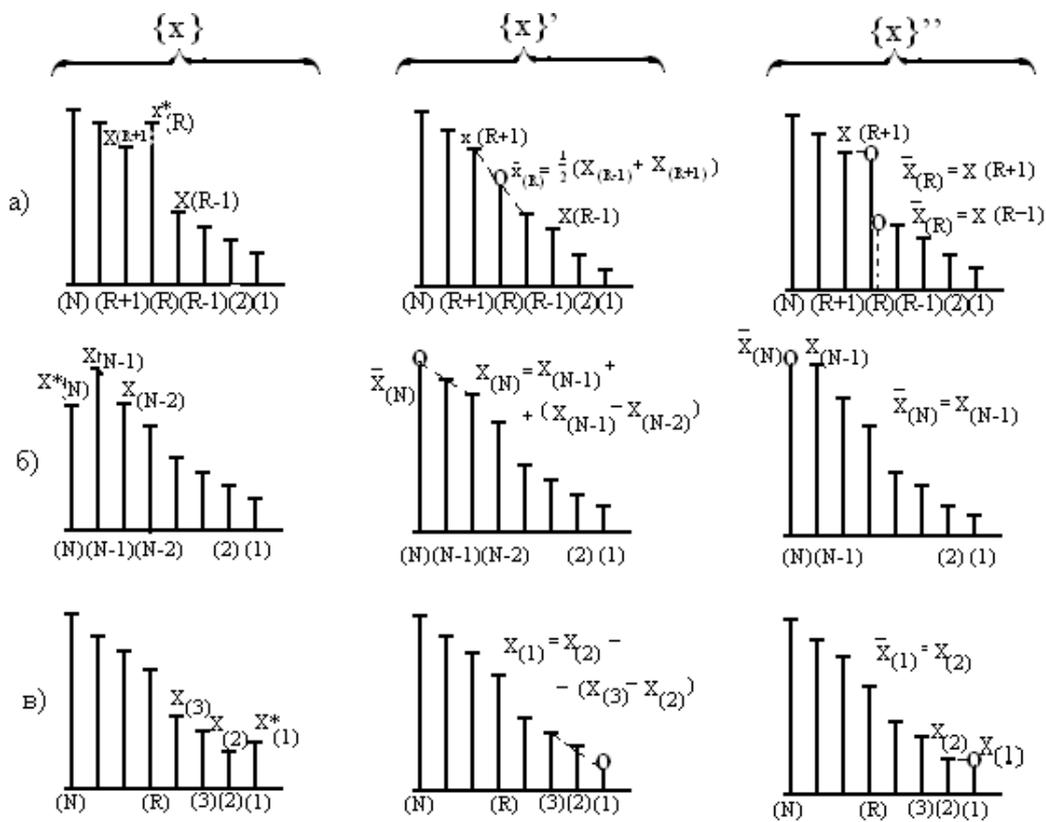


Рисунок 2. Коррекция ошибок на приемной стороне

Не приводя здесь количественные и качественные характеристики, отметим лишь, что апертурные методы сжатия данных оказались абсолютно непригодными для обработки при наличии в измерениях помех сбойного характера. При использовании же необратимых методов сжатия точность восстановления тренда практически не зависела от сбойных точек, тогда как зависимость от них оценок статистических характеристик стационарной составляющей была достаточно сильной.

Для исследования погрешности восстановления непрерывных реализаций с учетом искажений сжатых данных в канале связи моделировались процессы для разных способов передачи информации в стационарных и нестационарных каналах. При моделировании генерировались сигналы в виде суммы экспоненциального тренда и стационарной составляющей не дифференцируемого случайного процесса. Для сжатия использовались апертурные алгоритмы и алгоритмы, ограничивающие дисперсию ошибки восстановления. Учитывались искажения как информационной, так и случайной части сообщения. Адресная информация формировалась с каждым существенным отсчетом.

Результаты исследования

Полученные результаты свидетельствуют о том, что применение сложных алгоритмов сжатия целесообразно лишь для гладких, слабо зашумленных нестационарных сигналов и в стационарных каналах очень высокого качества (рис. 3).

При вероятностях ошибки приема символа сообщения $p > 10^{-2}$ и относительной мощности стационарной составляющей $h=1,0$ разница в погрешности восстановления нестационарных сообщений между простейшим алгоритмом ПО (предсказатель нулевого порядка) и достаточно сложным А0 – А1 (интерполяторы нулевого и первого порядка) становится пренебрежимо малой. Наличие искажений адресной информации в стационарном канале (рис. 4) почти на порядок повышает погрешность восстановления при апертуре $\epsilon > 3\%$, но мало влияет на нее при незначительных апертурах ($\epsilon < 1\%$).

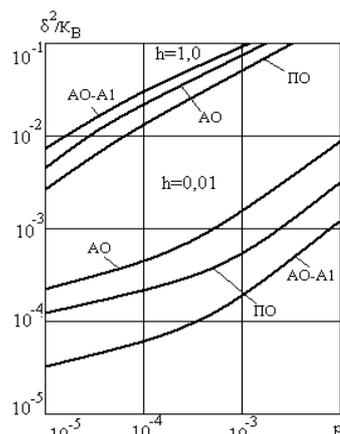


Рисунок 3. Погрешности восстановления для различных алгоритмов сжатия данных

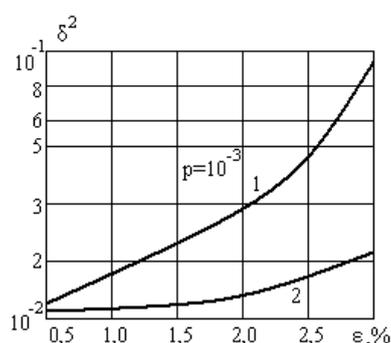


Рисунок 4. Погрешности восстановления нестационарных сообщений

Применение алгоритмов сжатия с контролем среднеквадратической погрешности восстановления (рис. 5) в нестационарном канале целесообразно лишь при очень малых апертурах (менее 1%) и малой относительной мощности стационарной составляющей сигнала ($h=0,1$). Для сильно зашумленных сигналов ($h=1,0$) погрешность восстановления практически не зависит от апертуры, а значит, и от коэффициента сжатия. При сжатии алгоритмом ПО эта тенденция проявляется еще резче.

На погрешность восстановления увеличение числа r проверочных символов корректирующего кода слабо влияет после $r=4$ (рис. 6) в плохом канале ($p=10^{-1}$).

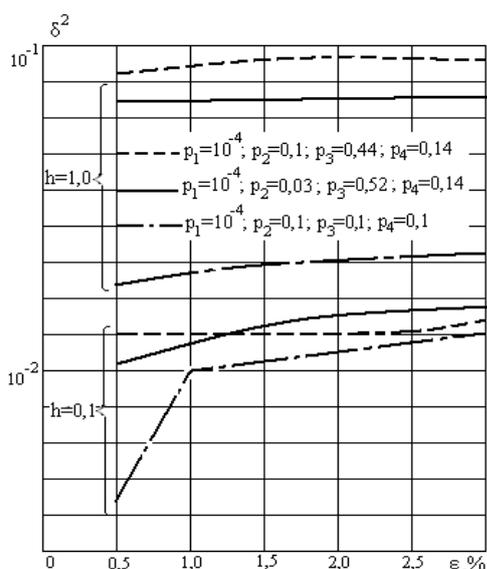


Рисунок 5. Погрешности восстановления в нестационарном канале

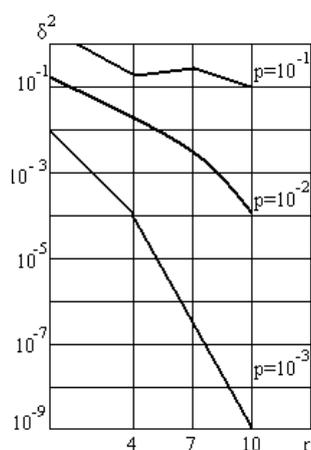


Рисунок 6. Погрешность восстановления при использовании корректирующего кода

В хорошем канале ($p < 10^{-4}$) погрешность оказывается приемлемой и без кодирования, поэтому можно рекомендовать применение для защиты адресной, временной и информационной частей сообщения коротких простейших циклических кодов Хэмминга. Этот вывод является следствием того, что группирование ошибок при передаче сжатых данных в каналах с памятью существенно (на 1-2 порядка) снижает погрешность восстановления сообщений по сравнению с передачей их в каналах с независимыми ошибками. В каналах с памятью исследовалось также влияние на погрешность восстановления нестационарных сигналов способов повышения достоверности передачи, таких, как использование обратного канала (с решающей и управляющей обратной связью) и отбраковка недостоверных измерений на приемной стороне системы.

Дискуссия

В литературных источниках проводились исследования погрешностей восстановления сжатых данных для стационарных каналов с нормальной помехой и стационарных передаваемых процессов [7,8]. При формировании служебной информации с каждым существенным отсчетом погрешность восстановления сигналов, вызванная искажениями служебной информации, оказалась на 2-3 порядка ниже, чем погрешность из-за искажения информационной части. Однако, реальные каналы связи часто оказываются нестационарными с наличием помех, коррелированных с полезным сигналом [9,10]. Такие каналы вносят существенные искажения при восстановлении сигналов на приемной стороне. Поэтому возникает необходимость обеспечения достоверности данных измерений на всех этапах преобразования данных, в том числе на этапе сжатия.

Как показали результаты имитационного моделирования, применение решающей обратной связи с переспросом ошибочно принятых сообщений уменьшает погрешность восстановления в среднем на 1 - 2 порядка при использовании простейшего кода с одним проверочным символом и на 2-3 порядка при 10 проверочных символов. При этом снижение коэффициента сжатия полосы частот прямого канала за счет переспросов не превышает 20%.

Однако применение решающей обратной связи в телеметрии имеет ограничения в связи с трудностью организации обратного канала и большими задержками в приеме ответов на значительных удалениях источника от приемника [11,12]. Поэтому были рассмотрены также возможности повышения точности восстановления при передаче сжатых данных по каналам с памятью путем отбраковки недостоверных измерений на приемной стороне. Изучалось влияние как параметрических методов (использование текущих разностей), так и непараметрических (с помощью статистики Диксона).

Результаты исследований свидетельствуют о том, что погрешность восстановления непрерывных процессов существенно зависит от выбора порога решающего правила, характера процесса и состояния канала передачи. Существует экстремум относительной погрешности восстановления (рис. 7), однако при гладких процессах ($h=0,1$) ошибка в выборе положения порога при недостатке априорной информации приводит к росту погрешности на 4-5 порядков по сравнению с минимальной (при использовании для отбраковки первых разностей).

Эффективность отбраковки для таких процессов также резко падает при средней длине пакетов ошибок $l_{cp} > 30$. Погрешность восстановления сильно зашумленных нестационарных сообщений (рис. 7, а) может быть снижена за счет исключения недостоверных данных примерно на порядок в каналах с большим уровнем помех ($p_e > 10^{-2}$, рис. 7, б); в каналах же высокого качества ($p_e < 10^{-3}$) этот выигрыш оказывается незначительным (рис. 7, в).

Лучшие результаты для обнаружения недостоверных измерений достигаются применением непараметрической статистики Диксона, которая за счет непараметричности обеспечивает примерно одинаковую эффективность и для гладких, и для зашумленных сигналов в каналах передачи разного качества.

Использование обратного канала с управляющей обратной связью показало, что ее применение позволяют существенно улучшить адаптацию системы к изменению помеховой обстановки в канале.

Таким образом, при передаче сжатых данных по стационарным каналам связи необходимости в помехоустойчивом кодировании не возникает даже при высоком уровне помех с вероятностью искажения одиночных символов сообщения порядка 10^{-2} .

И лишь в нестационарных каналах с памятью целесообразно применение коротких циклических кодов с 3-4 проверочными символами или сверточных кодов (при средней длине пакетов ошибочных измерений 2-20).

Более эффективным в этих случаях оказывается исключение недостоверных измерений с помощью непараметрических статистик вместе приема или использования обратного канала с решающей обратной связью.

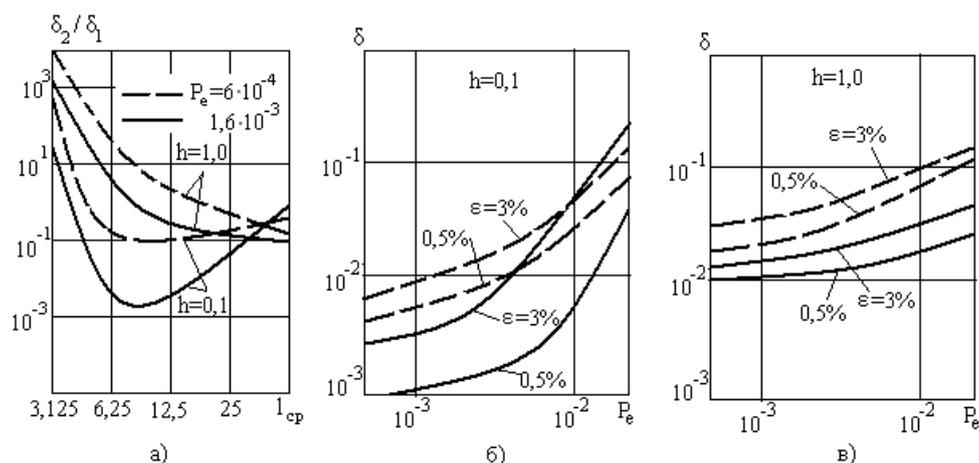


Рисунок 7. Эффективность отбраковки недостоверных данных

Заключение

Из проведенных исследований можно сделать следующие выводы.

1. При передаче сжатых данных по стационарным каналам связи необходимость в помехоустойчивом кодировании не возникает.
2. В нестационарных каналах с памятью целесообразно применение коротких циклических кодов с 3-4 проверочными символами или сверточных кодов.
3. Исключение недостоверных данных эффективно с помощью непараметрических статистик в месте приема или использование обратного канала с решающей обратной связью.

Список использованной литературы:

- 1 Yesmagambetov B.-B.S., Inkov A.M. Fast changing processes in radiotelemetry systems of space vehicles. *Journal of Systems Engineering and Electronics*. Vol. 26, No. 5. Beijing, October 2015, pp.941 – 945
- 2 Yesmagambetov B.-B.S., Ajmenov Zh., Inkov A. Ismailov S., Saribaev A. Statistical data processing in rocket-space technology. *Modern Applied Science*, Vol.9, №8. Canadian Center of Science and Education, 2015. P317-334.
- 3 Sprent P., Smeeton N.C. *Applied Nonparametric Statistical Methods*. Chapman and Hall/CRC, 2001, 462 pages.
- 4 Есмагамбетов Б.-Б.С. Статистическая обработка данных в радиотелеметрических системах. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*. №1.- М., 2015. С.13-21
- 5 Paul H. Kvat, Brani Vidakovic. *Nonparametric Statistics with Applications to Science and Engineering*. John Wiley & Sons, Inc, 2007
- 6 Gamiz M.L. et al *Applied Nonparametric Statistics in Reliability*. Springer, 2011, 230 pages.
- 7 Ю.Б. Ольховский, О.Н. Новоселов, А.П. Мановцев. *Сжатие данных при телеизмерениях*. – М.: Советское радио, 1971. – 303 с.
- 8 В.А. Свириденко. *Анализ систем со сжатием данных*. – М.: Связь, 1977. – 184 с.
- 9 А.Н. Дядюнов, Ю.А. Онищенко, А.И. Сенин. *Адаптивные системы сбора и передачи аналоговой информации*. – М.: Машиностроение, 1988. – 287 с.
- 10 Назаров А.В., Козырев Г.И., Шитов И.В. и др. *Современная телеметрия в теории и на практике. Учебный курс*. – Санкт-Петербург: Наука и техника, 2007. – 667 с.
- 11 Б.Я.Авдеев, Е.М. Антонюк, С.Н. Долинов и др.. *Адаптивные телеизмерительные системы*. – Л.: Энергоиздат, 1981. – 248 с.
- 12 Победоносцев В.А. *Основы информметрии*. – М.: «Радио и связь», 2000. – 192 с.

References:

- 1 Yesmagambetov B.-B.S., Inkov A.M. (2015) Fast changing processes in radiotelemetry systems of space vehicles. *Journal of Systems Engineering and Electronics*. Vol. 26, No. 5. Beijing, October, pp.941 – 945
- 2 Yesmagambetov B.-B.S., Ajmenov Zh., Inkov A. Ismailov S., Saribaev A. (2015) Statistical data processing in rocket-space technology. *Modern Applied Science*, Vol.9, №8. Canadian Center of Science and Education, P317-334.
- 3 Sprent P., Smeeton N.C. *Applied Nonparametric Statistical Methods*. Chapman and Hall/CRC, 2001, 462 pages.

- 4 Yesmagambetov B.-B.S. (2015) *Statisticheskaja obrabotka dannyh v radiotelemetricheskih sistemah*. [Statistical processing of data in radiotelemetry systems]. *Vestnik MGTU im. N.E. Baumana. Ser. Priborostroenie*. No. 1. M., 13-21). (In Russian).
- 5 Paul H. Kvam, Brani Vidakovic. (2007) *Nonparametric Statistics with Applications to Science and Engineering*. John Wiley & Sons, Inc.
- 6 Gamiz M.L. (2011) *et al Applied Nonparametric Statistics in Reliability*. Springer, 230.
- 7 Yu. B. Olhovski, O.N. Novoselov, A.P. Manovtcev. (1971) *Szhatie dannyh pri teleizmereniih* [Data Compression for Tele-Measurements]. *Sovetskoe radio*. M., 303. (in Russian).
- 8 V.A. Sviridenko. *Analiz system so szhatiem dannyh* (1977) [Data Compression System Analysis]. *Sviaz*. – M., 184. (In Russian).
- 9 A.N. Diadyunov, Yu. A. Onishenko, A.I. Senin. (1988) *Adaptivnye sistemy sbora i peredachi analogovoi informatsii* [Adaptive analog data collection and transmission systems]. *Mashinostroenie*. M., 287. (In Russian).
- 10 Nazarov A.V., Kozyrev G.I., Shitov I.V. I dr. (2007) *Sovremennaja telemetriia v teorii i na praktike. Uchebnyi kurs* [Modern telemetry in theory and practice. Training course]. *Nauka i tehnika*. St. Petersburg, 667. (In Russian).
- 11 B. Ia. Avdeev, E.M. Antonyuk, S.N. Dolinov I dr. (1981) *Adaptivnye teleizmeritelnye sistemy* [Adaptive Telemetry Systems]. *Energoizdat*. L, 248. (In Russian).
- 12 Pobedonostcev V.A. *Osnovy informmetrii* (2000) [Fundamentals of Information Metrics]. *Radio i sviaz*. M., 192. (In Russian).