АЛГОРИТМЫ ОПТИМИЗАЦИИ ПРОПУСКНОЙ СПОСОБНОСТИ КАНАЛА СВЯЗИ ДЛЯ ЗАЩИЩЕННЫХ СИСТЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Кутья А.С.

Дубровский В.В. – кандидат физико-математических наук, доцент

Тема работы: «Алгоритмы оптимизации пропускной способности канала связи для защищенных систем передачи информации».

Актуальность темы исследования обусловлена тем, что важным элементом любого канала передачи данных, независимо от ее сложности и назначения, являются программные и программно-аппаратные средства кодирования.

Цель работы состоит в оценке пропускной способности канала связи для защищенных систем передачи информации. Необходимо решить проблему информационной безопасности инфокоммуникационной системы, функционирующей в условиях действия аддитивных и мультипликативных помех.

В связи с интенсивным развитием цифровых систем передачи и обработки информации актуальной задачей является обеспечение высокой ее достоверности или минимальной вероятности ошибочного приема. Эффективным способом решения этой задачи является применение помехоустойчивого или корректирующего кодирования информации. Следовательно, выбор помехоустойчивого кода, метода кодирования и алгоритма декодирования информации является актуальной задачей теории и практики помехоустойчивого кодирования.

Для передачи данных по моделируемому защищенному каналу используется модифицированный код Хэмминга. Параметры кодов Хемминга: $n = 2^m - 1,\ m > 1;\quad k = n - m;\quad d = 3;$ проверочная матрица $H = (1, \alpha, \alpha^2, ..., \alpha^{n-1}),$ где α – примитивный элемент поля Галуа $GF(2^m)$ [1].

Столбцы проверочной матрицы являются элементов поля $GF(2^m)$, то есть векторами из P_n в базисе $1, \alpha, \alpha^2, ..., \alpha^n$ для примитивного элемента α поля $GF(2^m)$ [2].

Расстояние Хэмминга между двумя векторами $\bar{x}, \bar{y} \in P_n$ – это число $dist(\bar{x}, \bar{y})$ несовпадающих координат данных векторов. Весом $wt(\bar{x})$ векторов из пространства $\bar{x} \in P_n$ называют число ненулевых координат этих векторов [2].

Предлагаемый алгоритм обладает следующими свойствами:

1) Все одиночные битовые ошибки могут быть исправлены;
2) Все двойные битовые ошибки могут быть обнаружены;
3) Все соседние битовые двойные ошибки могут быть исправлена
4) Вероятность неверного исправления для несмежных двойных ошибок снижена.

Характеристики линейного блочного кода полностью определяется его H-матрицей. Для того, чтобы обнаружить все одиночные битовые ошибки, соответствующие синдромы ошибок должен быть уникальными. Следует учитывать, что синдром для однобитовой ошибки в бите с позицией p совпадает с p-й столбцом H-матрицы. Для того, чтобы однозначно идентифицировать все одиночные битовые ошибки, все столбцы H-матрицы должны быть уникальными [3].

Для того, чтобы обнаружить все двойные ошибки в битах, соответствующий синдромы должны отличаться от всех синдромов однобитовых ошибок.

Синдром для двойной битовой ошибки определяется операцией исключающее ИЛИ (XOR) соответствующих столбцов H-матрицы. Так что не может быть 3-цикла в H-матрицах. *k*-цикл относится к набору k линейно зависимых столбцов матрицы проверки на четность, то есть, когда проведены все операции XOR, в результате имеется полностью нулевой столбец. Для исправления всех смежных двойных битовых ошибок, синдромы близлежащих двойных битовых ошибок должны быть отличны друг от друга, а также отличается от синдромов всех однобитовых ошибок [4].

Определим условия, которым должны удовлетворять H-матрицы для предлагаемого кода:

1) Не все столбцы нулевые;
2) Все столбцы являются различными;
3) Отсутствует линейная зависимость, включающая 3 или меньше столбцов т.е. отсутствуют 2-циклы, 3 циклы допускаются [5].
4) Отсутствует линейная зависимость столбцов с участием C_i, C_j, C_k, C_m, где $m > k > j > i$, такие, что $j = i + 1$ и $m = k + 1$.
5) Кроме того, код пытается минимизировать количество 4-циклов с участием C_i, C_j, C_k, C_m, где $m > k > j > i$, такие, что $j = i + 1$ и $m = k + 1$.

Первый используемый набор данных - передача 120 МБ информации при нормальных условиях эксплуатации сети.

Выборка данных проводилась каждые 0,1 секунды, и было зарегистрировано 200 пропусков.

Данные были разделены на наборы данных обучения и тестирования. Из-за ограниченности имеющихся данных, одна седьмая данных была сохранена в качестве набора тестов, а остальные представлены для обучения.

Для эксперимента с данными была реализована нейронная сеть - генетический алгоритм (NN-GA) с использованием сети кодеров, обученной 4 скрытыми узлами в течение 200 тренировочных этапов [6].

Результаты тестирования (рисунок 1) показывают, что алгоритм не позволил сделать прогноз для столбца 1 в этом наборе данных. Причина состоит в том, что для данного метода, чтобы сделать прогноз, матрица прогнозирования должна быть положительной. Основной причиной этого является то, что одна переменная линейно зависит от другой переменной. Эта линейная зависимость может иногда существовать не между самими переменными, а между элементами моментов, такие как среднее, дисперсии, ковариации и корреляции. Другие причины этой причины включают ошибки при чтении данных, начальные значения и многое другое. Эта проблема может быть решена путем удаления переменных, которые линейно зависят друг от друга, или путем использования главных компонентов для замены набора коллинеарных переменных на ортогональные компоненты. Для остальных наборов в других столбцах результаты показывают удовлетворительный уровень восстановления пропущенных данных.

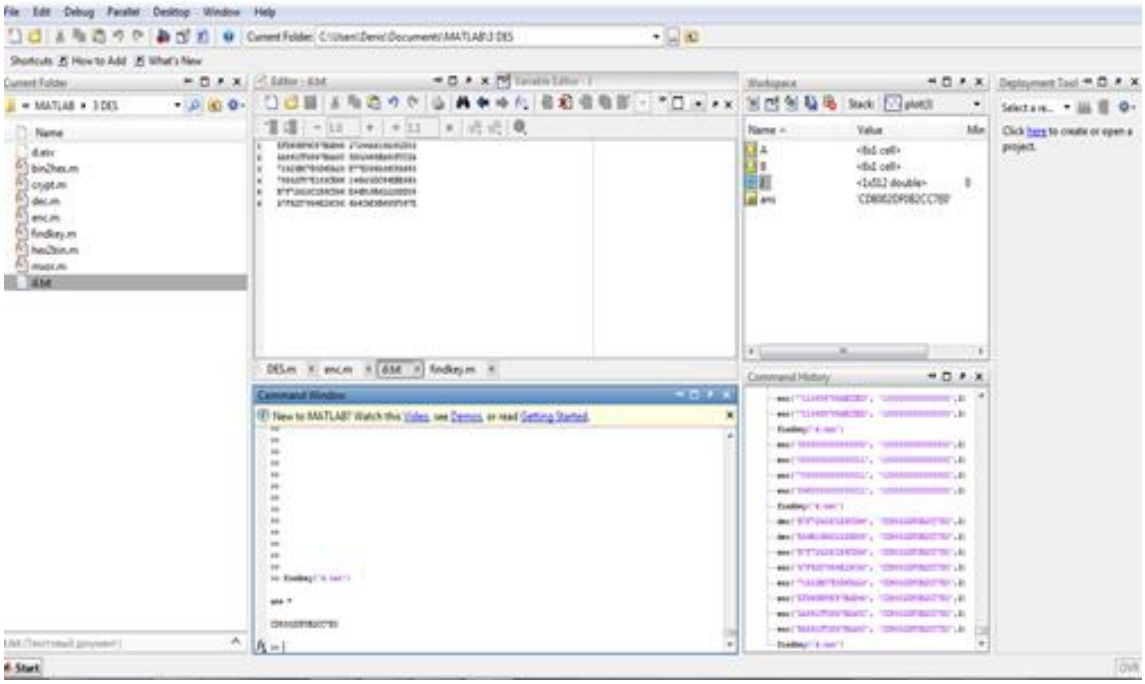

Рис. 1 – Результаты тестирования

Нейробайесовские модели являются концептуально естественным подходом для применения в IT-сфере. Современные вероятностные языковые среды программирования для байесовских вычислений еще больше упростили его применение, предоставив интерфейсы для определения потенциально очень сложных моделей даже для неспециалистов. В этом исследовании описаны базовые теоретические основы, необходимые для реализации байесовского моделирования с упором на приложения в информационной безопасности. Тем не менее, необходимы дальнейшие исследования улучшенных и более быстрых методов байесовских вычислений для больших данных. Байесовское моделирование требует значительного количества предположений о составлении порождающих моделей и уточнении исходных предположений.

Список использованных источников:

1. Guipta A., Lams M.S. Estimation Missings Value using Neural Network // J. of Operationals Research Society. – 2016. – Vol. 48. – № 1. – С. 329–339.
2. Nelamondo F.V., Mohammed S., Marwalas T. Missing Datas: A comparisons of neural networks and expectation maximizations technique // Currents Science. – 2017. – Vol. 95. – № 12. – С.467–473.
3. Карловf И.А. Методы определения пропускной способности с использованием инструментария Data Minings // Вестник Сиб. гос. аэрокосмического ун-та им. Ак. М.Ф. Решетнев. – 2015. – № 7(41) – С.39–43.
4. Карлова И.А., Кашур В.Д. Подход к построению гибридных моделей для оценки значения пропущенного элемента в массиве данных //: Матер. XX Всеросс. семинара. – 2012. – С. 174–179.
5. Halkidi M., Batistakis Y., Vazirgiannis M. On Clustering Validation Techniques // J. of Intelligent Information Systems. – 2013. – № 17:2/3.– С.107–145.
6. An Introductions to computing with neural net, Richard P. Lipman, IEE ASP Magazines, April 2017, pages 2-22.
7. A Neural Networks Approaches Toward Intrusion Detections, Kevins L. Foxer, Rondal R. Hennings, Jonatan H. Reeds, Richard P. Sitnonians, Harris Corporations, Government Informations System Divisions, P.O. Box 98000, Melbourne, FL 32902, July 2010.
8. Univariate Economic Time Series Forecasting by Connexionist Methods, A. Varfis and C. Versino, Proceedingsof the InternationalNeuralNetworksConference, Paris, 2010, pages 342-345.