ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

Федеральное государственное образовательное бюджетное учреждение высшего образования

«ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ И ИНФОРМАТИКИ»

Кафедра «Информатики и вычислительной техники»

Сдана на проверку Допустить к защите

«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г. «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.

Защищена с оценкой \_\_\_\_\_ «\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.

**ОБРАБОТКА ДАННЫХ И**

**МОДЕЛИРОВАНИЕ**

**В МАТЕМАТИЧЕСКИХ ПАКЕТАХ**

Пояснительная записка

на \_\_\_\_\_\_\_\_\_ листах

Студент (ка) группы\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Руководитель\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

№ зачётной книжки\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Самара 20\_\_г.

# **Рецензия**

**Оглавление**

[**Рецензия** 2](#_Toc450827170)

[**Оглавление** 3](#_Toc450827171)

[**Цели и задачи** 4](#_Toc450827172)

[**Постановка задачи курсовой работы** 5](#_Toc450827173)

[**Выбор исходных данных** 6](#_Toc450827174)

[**Оценка основных показателей информации** 7](#_Toc450827175)

[**Шифрование заданной фразы** 11](#_Toc450827176)

[**Математическое моделирование** 22](#_Toc450827177)

[**Выводы** 27](#_Toc450827178)

[**Список используемой литературы** 28](#_Toc450827179)

# **Цели и задачи**

Целью курсовой работы является развитие практических навыков работы в системах *Mathcad* и *Matlab + Simulink,* как наиболее широко используемых в области телекоммуникаций.

Задачей курсовой работы является цифровая обработка данных с использованием языков программирования в математических средах и моделирование соответствующих устройств.

## **Постановка задачи курсовой работы**

1. Для заданной фразы (сообщения S) оценить основные показатели информации:

* информационный объём одного символа;
* объём данных сообщения;
* количество информации;
* коэффициент информативности;
* вероятности появления символов;
* энтропию сообщения;
* избыточность сообщения.

Целью данного пункта задания является осмысление показателей информации и освоение операций над векторами данных в математических средах.

2. В системе *Matlab + Simulink* разработать программы шифрования и расшифрования заданной фразы по заданному алгоритму и провести моделирование этих процессов.

Целью данного пункта задания является освоение приёмов программирования и создания пользовательских блоков в системе *Matlab + Simulink* на примере одного из методов защиты информации.

3. Для совокупности кодов любых трёх смежных слов заданной фразы в системе *Mathcad* с помощью регрессионного анализа выполнить математическое моделирование выборки. При этом аналитическое выражение (математическая модель) должно наиболее достоверно описывать данную выборку.

Целью данного пункта является получение представления о математическом моделировании относительно интерполяции данных, имеющей большое значение в задачах сжатия цифровых видео- и звуковых сигналов.

# **Выбор исходных данных**

 С учётом номера в списке группе, был выбран седьмой вариант. Фраза для шифрования для седьмого варианта:

«Лучший дар детям – не собственность, а воспитание в них работоспособности и совести».

# **Оценка основных показателей информации**

 Для заданной строки S были рассчитаны основные показатели информации: информационный объём одного символа, объём данных, количество информации, коэффициент информативности, вероятности появления символов в сообщении, энтропия и избыточность сообщения.

 Затем для этой фразы были вычислены статистические показатели вероятностей появления символов и построена гистограмма частости распределения символов в строке.



Рисунок 1. Расчёт основных показателей информации

 На рисунке 1 показано, как с помощью средств математического пакета *Mathcad* были рассчитаны количество информации на 1 двоичный символ, объём данных строки S, количество информации в строки и её коэффициент информативности согласно известным из курса информатики формулам.

 На рисунках 2 и 3 представлен расчёт энтропии и избыточности сообщения заданной строки S.



Рисунок 2. Расчёт энтропии и избыточности сообщения



Рисунок 3. Расчёт энтропии и избыточности сообщения

На рисунке 4 представлена диаграмма частости распределения символов в строке, построенная средствами математической среды Mathcad на основании проведённых расчётов.



Рисунок 4. Диаграмма частости распределения символов в строке

 Вывод: Математическая среда *Mathcad* является простым и понятным инструментом даже для начинающих пользователей. Набор инструментов, предоставляемый данной программой, позволяет производить как простые арифметические подсчёты, так и сложные технические расчёты с использованием векторов и матриц. В дополнение к этому в *Mathcad* присутствует возможность построения графиков и диаграмм, что значительно улучшает визуальное восприятие результатов расчётов.

Данный математический пакет позволил довольно просто и быстро произвести расчёты основных показателей информационного сообщения. В результате мной были получены навыки работы с векторами и графиками в данной среде, а также поняты основные информационные показатели, что удовлетворяет целям и задачам курсовой работы.

# **Шифрование заданной фразы**

 С помощью программной среды *Matlab + Simulink* было произведено шифрование заданной фразы S методом, выбранным в соответствии с вариантом. Метод шифрования варианта 7:

1. строковая фраза заменяется кодами символов;
2. код смещается на N+2 позиции вправо и к вновь полученному коду добавляется циклическая гамма: 1 4 2 8 5 6 9.

Реализация шифрования фразы с помощью прямых вычислений:

>> S='Лучший дар детям – не собственность, а воспитание в них работоспособности и совести.';

>> G='142856914285691428569142856914285691428569142856914285691428569142856914285691428569';

>> C=double(S) %Преобразование фразы в код

C =

 Columns 1 through 8

 1051 1091 1095 1096 1080 1081 32 1076

 Columns 9 through 16

 1072 1088 32 1076 1077 1090 1103 1084

 Columns 17 through 24

 32 8211 32 1085 1077 32 1089 1086

 Columns 25 through 32

 1073 1089 1090 1074 1077 1085 1085 1086

 Columns 33 through 40

 1089 1090 1100 44 32 1072 32 1074

 Columns 41 through 48

 1086 1089 1087 1080 1090 1072 1085 1080

 Columns 49 through 56

 1077 32 1074 32 1085 1080 1093 32

 Columns 57 through 64

 1088 1072 1073 1086 1090 1086 1089 1087

 Columns 65 through 72

 1086 1089 1086 1073 1085 1086 1089 1090

 Columns 73 through 80

 1080 32 1080 32 1089 1086 1074 1077

 Columns 81 through 84

 1089 1090 1080 46

>> CG=double(G) %Преобразование гаммы в код

CG =

 Columns 1 through 9

 49 52 50 56 53 54 57 49 52

 Columns 10 through 18

 50 56 53 54 57 49 52 50 56

 Columns 19 through 27

 53 54 57 49 52 50 56 53 54

 Columns 28 through 36

 57 49 52 50 56 53 54 57 49

 Columns 37 through 45

 52 50 56 53 54 57 49 52 50

 Columns 46 through 54

 56 53 54 57 49 52 50 56 53

 Columns 55 through 63

 54 57 49 52 50 56 53 54 57

 Columns 64 through 72

 49 52 50 56 53 54 57 49 52

 Columns 73 through 81

 50 56 53 54 57 49 52 50 56

 Columns 82 through 84

 53 54 57

>> %Формирование шифрограммы кодов строки S

>> N=1104;

>> CK=mod(C+N+2+CG,1104) %Замена кодов строки в другой код

CK =

 Columns 1 through 8

 1102 41 43 50 31 33 91 23

 Columns 9 through 16

 22 36 90 27 29 45 50 34

 Columns 17 through 24

 84 541 87 37 32 83 39 34

 Columns 25 through 32

 27 40 42 29 24 35 33 40

 Columns 33 through 40

 40 42 55 95 86 20 90 25

 Columns 41 through 48

 38 44 34 30 38 26 36 32

 Columns 49 through 56

 32 83 24 84 39 31 45 91

 Columns 57 through 64

 35 22 21 40 41 38 44 34

 Columns 65 through 72

 36 37 40 24 37 41 36 40

 Columns 73 through 80

 28 90 31 88 44 33 24 25

 Columns 81 through 84

 43 41 32 105

>> %Получение шифрограммы символов

>> CH=char(CK)

CH =

ю)+2¬

Рисунок 5. Программа шифрования в среде Matlab



Рисунок 6. Программа дешифровки в среде Matlab

 Результат использования программ в виде кода:

>> S='Лучший дар детям – не собственность, а воспитание в них работоспособности и совести.';

>> G='142856914285691428569142856914285691428569142856914285691428569142856914285691428569';

>> CK=SC(S,G);

>> CH=char(CK)

CH =

ю)+2¬

Рисунок 7. Модель работы программ шифрования и дешифровки



Рисунок 8. Коды строки S



Рисунок 9. Коды гаммы



Рисунок 10. Коды шифрограммы



Рисунок 11. Коды шифрограммы при дешифровке



Рисунок 12. Коды строки S после дешифровки

 Код результатов выполнения программ:

>> S='Лучший дар детям – не собственность, а воспитание в них работоспособности и совести.';

>> C=double(S);

>> G='142856914285691428569142856914285691428569142856914285691428569142856914285691428569';

>> CG=double(G);

>> CK=SCS(C,CG);

>> char(CK)

ans =

ю)+2¬

Рисунок 13. Решение задачи в системе Mathcad

 Как видно, *Mathcad* справился с данной задачей лучше, потому что *Matlab* не смог дешифровать знак “-“.

 Для решения данной задачи подходят оба математических пакета, однако ­*Mathcad* справился лучше.

 В процессе выполнения задачи возникла проблема переименования блоков в новых версиях системы ­*Matlab*. В новых версиях блоки пользовательских функций устроены более сложно, поэтому пришлось искать им замену в виде блоков интерпретируемых функций *Matlab*.

 Данная задача помогла мне развить навыки программирования в средах математических пакетов для выполнения задач разной структуры и сложности.

# **Математическое моделирование**

 Средствами математического пакета *Mathcad* была рассмотрена возможность проведения регрессионного анализа для восстановления фразы из её кодов. В процессе выполнения задачи были использованы три закона для последующего регрессионного анализа и выбора наиболее подходящего для дешифровки закона.

 На рисунке 14 представлены модули формирования вектора выборки данных для фразы, созданные средствами программирования *Mathcad.*



Рисунок 14. Формирование вектора выборки данных

 Для анализа была выбрана самая короткая последовательность слов, состоящая из 12 элементов. Затем она была записана в строковую переменную без пробелов и заглавных букв. Это сделано для того, чтобы при анализе не возникло проблем с описанием резких скачков и падений значений ввиду расположения кода символа в таблице кодов Unicode.

 На рисунке 15 продемонстрирован анализ методом линейной регрессии.



Рисунок 15. Линейная регрессия

 На рисунке видно, что линейная регрессия и её аналитическое выражение не подходят для восстановления фразы. Это можно утверждать как на основании низкого коэффициента детерминации, так и на основании графика, построенного согласно вектору выборки (синяя сплошная линия) и вектору результатов анализа (чёрная пунктирная линия).

 На рисунках 16-17 представлен анализ фразы с помощью полиноминальной функции.



Рисунок 16. Полиноминальная функция



Рисунок 17. Вывод и аналитическое выражение полиноминальной функции

 Как видно, полиноминальная функция даёт коэффициент детерминации, равный 1. График данного аналитического выражения также совпадает с графиком вектора выборки.

 На рисунке 18 представлена линейная регрессия общего вида.



Рисунок 18. Линейная регрессия общего вида

 На данный закон было потрачено наибольшее количество времени, потому что для неё необходимо подобрать комбинацию функций, описывающих график. Несмотря на это, мне так и не удалось подобрать точную комбинацию функций. Вывод: Линейная регрессия общего вида не подходит для данной фразы.

 Исходя из полученных результатов, мной был выбран закон полиноминальной функции для дешифровки сообщения, как наиболее правильный.

 На рисунке 19 представлен процесс дешифровки и конечный результат.



Рисунок 19. Дешифровка и сравнение фраз

 Даже учитывая то, что коэффициент детерминации для полиноминальной функции был равен единице, дешифровка фразы произошла с ошибками. Лишь 8 символов удалось декодировать однозначно. Из этого следует, что регрессионный анализ для текстовых данных не может дать адекватного результата.

 В процессе работы мною были приобретены навыки регрессионного анализа средствами математического пакета *Mathcad,* а также получены знания о таких законах регрессионного анализа, как линейная регрессия, полиноминальная регрессия и линейная регрессия общего вида.

# **Выводы**

 Курсовая работа по курсу «Информатика» охватывает основные аспекты используемых информационных технологий, связанных с направлением «Инфокоммуникационные технологии и системы связи». В процессе работы мной были приобретены навыки вычисления основных информационных показателей для текстовых сообщений с использованием средств современных математических пакетов *Mathcad* и *Matlab + Simulink,* навыки регрессионного анализа и моделирования.

 В процессе решения задач курсовой работы мной были приобретены практические навыки программирования для цифровой обработки данных с использованием современных программируемых математических сред.

# **Список используемой литературы**