

Министерство образования и науки Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего профессионального образования  
«Санкт-Петербургский политехнический университет Петра Великого»  
Институт компьютерных наук и технологий

**Кафедра компьютерных систем и программных технологий**

## **ВЫПУСКНАЯ РАБОТА БАКАЛАВРА**

**Разработка помехоустойчивого кодека с применением технологий  
дизассемблирования и реверс-инжиниринга**

направление: 09.03.01 - Информатика и вычислительная техника

Выполнил:  
Чеботарев Георгий (Михайлович)  
Подпись \_\_\_\_\_  
Руководитель:      доцент,      к.т.н.,  
Богач Наталья Владимировна  
Подпись \_\_\_\_\_

Санкт-Петербург

2016



## **Реферат**

ДИЗАССЕМБЛИРОВАНИЕ ПРОШИВКИ, РЕВЕРС-ИНЖИНИРИНГ, ПОМЕХОУСТОЙЧИВЫЙ КОДЕК, ЯЗЫК ПРОГРАММИРОВАНИЯ СИ/C++.

В выпускной работе проводится дизассемблирование прошивки устаревшего устройства, с целью изучения и модернизации алгоритмов помехоустойчивого кодирования, для последующего использования в новом устройстве. В работе описаны наиболее популярные из существующих дизассемблеров. Дано краткое описание используемого микропроцессора TMS320C54. Приведена алгоритма Витерби. Описан жизненный цикл разработки

.....Полный шлак еще думать

## Содержание

Реферат .....	3
Введение .....	5
1. Особенности микропроцессора TMS320C54 .....	6
2. Описание алгоритма Витерби .....	7
3. Жизненный цикл разработки помехоустойчивого кодека .....	7
3.1. Дизассемблирование.....	9
3.1.1. Выбор среды дизассемблирования .....	9
3.1.2. Процесс дизассемблирования прошивки .....	11
3.1.3. Выбор симулятора микропроцессора.....	15
3.1.4. Настройка среды симулятора.....	15
3.2. Анализ полученного кода .....	18
3.2.1. Поиск функции кодека.....	19
3.2.2. Подтвердить работоспособность кодека.....	20
3.2.3. Анализ структуры кодека .....	21
3.3. Разработка модулей кодека на языке Си .....	23
3.3.1. Кодер.....	24
3.3.2. Модули Чередования .....	25
3.3.3. Декодер.....	26
3.4. Тестирование реализации на языке Си.....	27
3.5. Модернизация кодека .....	29
3.5.1. Кодер.....	30
3.5.2. Модули Чередования .....	30
3.5.3. Декодер.....	30

3.6. Тестирование декодирующих способностей всех трех режимов .....	32
Заключение.....	36
Список использованных источников .....	38

## **Список рисунков**

## **Введение**

Помехоустойчивое кодирование данных является повседневной задачей. Данные кодируется для любого типа каналов и видов передач. Без такого рода защиты от шума в канале передачи, отправка даже небольших объемов данных, на незначительное расстояние, представлялась бы невозможной задачей. Количество повторных передач возросло бы в сотни раз, что конечно бы негативно отразилось на организации работы всех систем передачи данных.

Цель помехоустойчивого кодирования – это минимизация искажения данных, в следствии воздействия на них шумовых эффектов, в канале передачи. Для этого применяются различные приемы: кодирование данных, чередование закодированного сообщения, повторное кодирование и т.д.

При кодировании данные представляются специальным образом. Если после передачи данных по каналу, в закодированном сообщении возникнут изменения, при декодировании это можно будет обнаружить и исправить.

Одним из самых распространённых алгоритмов декодирования является алгоритм Витерби, обладающий высокими восстанавливающими показателями, при относительно низких временных затратах.

## **Постановка задачи**

Необходимо провести обратную разработку кодека, скрытого в прошивке устройства, базирующегося на основе микропроцессора TMS320C54.

Необходимо дизассемблировать прошивку, определить границы кодека, проанализировать алгоритмы функций кодека и реализовать их на языке Си. Разработанный кодек должен поддерживать протокол работы кодека, скрытого в прошивке. Необходимо разработать и внедрить промежуточный режим кодирования (декодирования). Конечная версия кодека должна быть оформлена в виде библиотеки. Взаимодействие с памятью должно осуществляться посредством передачи указателей на участки одномерного массива внутрь функций кодека. По завершению разработки, необходимо провести сравнительный анализ всех режимов кодирования.

## **1. Особенности микропроцессора TMS320C54**

Разработка базируется на основе прошивки, изъятой из микропроцессора TMS320C54x, разработанного компанией Texas Instrument. Для понимания кода, полученного после дизассемблирования, необходимо понимание внутренней организации и особенностей этого микропроцессора. Семейство микропроцессоров TMS320C54x ориентировано на быструю обработку цифровых сигналов. К их архитектурным особенностям можно отнести следующие:

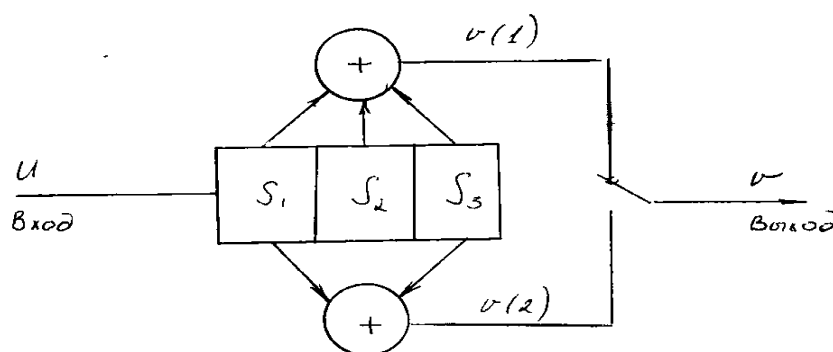
- Модифицированная Гарвардская архитектура с отдельными 16-битными шинами памяти: одна под программу, три под данные.
- Высокоскоростной АЛУ, обладающий высоким значением распараллеливания.
- Память и дополнительные периферийные устройства на чипе
- Наличие узкоспециализированных инструкций (например, для реализации «Витерби-бабочки»).
- Два 40-битных аккумулятора и восемь 16-битных регистров.
- Обладает 7 видами адресации.

Совокупность этих отличительных характеристик позволяет выполнять кодирование данных максимально эффективно. Более подробное описание микропроцессора можно прочитать на сайте производителя [1].

## 2. Описание алгоритма Витерби

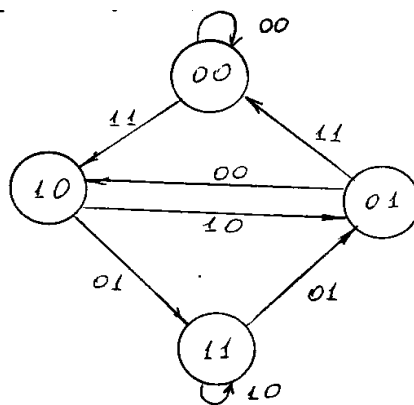
**Алгоритм Витерби** — это алгоритм динамического программирования, выполняющий поиск наиболее подходящего списка состояний и получающий наиболее вероятную последовательность произошедших событий. Алгоритм получил своё название в честь создателя — американского инженера итальянского происхождения Эндрю Витерби [2].

Алгоритм декодирует последовательность, полученную в результате работы сверточного кодера. Рассмотрим работу кодера на несложном примере (см. Рис. 1).



**Рис. 1. Пример сверточного кодера**

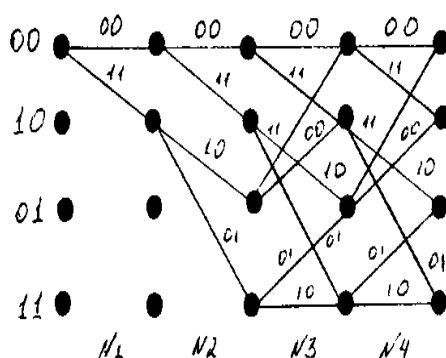
В зависимости от поступившего бита и линии задержки, создается закодированная пара бит, которая поступает на выход декодера. Для понимания работы алгоритма Витерби, сверточный кодер удобно будет представить в виде конечного автомата. (см. пример на Рис. 2).



**Рис. 2. Пример конечного автомата кодера**

В вершинах графа указаны состояния кодера, дуги — это значения выходов с кодера. В зависимости от того, в каком состоянии кодер находится, и того, что было принято на вход — кодер переходит из одного состояния в другое. Изначально кодер находится в состоянии 00, и если на вход поступит 1, кодер перейдет в состояние 10, а на выходе кодера будет 11.

Развертка диаграммы состояний во времени образует решетчатую диаграмму (см. Рис. 3).



**Рис. 3. Пример решетки Витерби**

Каждое состояние решетки соответствует состоянию конечного автомата. Ширина решетки равна количеству состояний в конечном автомате. Движение по решетке осуществляется слева направо. Вход в решетку, как и в графе, находится в состоянии 00. В зависимости от того, какая последовательность принимается на вход декодера, осуществляются переходы из одного состояния в другое. В случае же, если таких переходов из состояния



нет, например, из текущего состояния, в графе предусмотрены переходы с весами 10 и 01, а на вход была принята последовательность 00, каждый из маршрутов, берущий начала из данного состояния, получает штраф, за один не соответствующий бит. При работе алгоритма генерируется N-ое количество путей. Для сокращения вычислительных ресурсов, отслеживается только K путей из всех возможных (для данного примера  $K = 4$ ), остальные маршруты отбрасываются. В решетке остаются пути с наименьшим количеством штрафов, остальные отбрасываются.

По того, как вся решетка была построена, выбирается лучший путь и выполняется обратная проходка. При этом происходит восстановление и декодирование искомого сообщения.

### **3. Жизненный цикл разработки помехоустойчивого кодека**

#### **3.1. Дизассемблирование**

Дизассемблирование является ключом к исходному коду любой программы, будь это компьютерный вирус или любой другой бинарный файл. Транслирование исходного файла на язык ассемблер позволяет увидеть и изучить алгоритмы работы программы. В связи с тем, что технология применяется не повсеместно – качественных сред дизассемблирования немного.

##### **3.1.1. Выбор среды дизассемблирования**

К наиболее известным и активно используемым дизассемблерам можно отнести: Sourcer, Hiew и IDA Pro. Каждый из них доказал свое право на существования, однако использоваться будет IDA Pro. Этот выбор обуславливается положительными и отрицательными сторонами всех трех дизассемблеров:

- *Sourcer* – простейший дизассемблер. Позволяет конвертировать исходный файл в ассемблерный код, однако, примерно в 30% случаев требуется вмешательство программиста для исправления ошибок

дизассемблирования, что усложняет отладку большой программы. Невозможно добавлять комментарии или изменять название функций – это усложняет понимание кода и сказывается на скорости и разработки.

- *Hiew* - редактор двоичных файлов, ориентированный на работу с кодом. Имеет встроенный дизассемблер для x86, x86-64 и ARM, ассемблер для x86, x86-64 [3]. Стоит так же отметить, что данный редактор распространяется бесплатно. Однако с момента создания, начале 90-х, в Hiew серьезных изменений не вносили, программа не удобна при работе: полное отсутствие интерактивности и командный режим. Количество поддерживаемых форматов входных файлов и процессоров жестко ограничено. Очевидно, что недостатки превышают достоинства. Данный редактор подойдет для просмотра двоичного кода, но для дизассемблирования прошивки.
- *IDA Pro* (англ. Interactive DisAssembler) — интерактивный дизассемблер, отличается исключительной гибкостью, наличием встроенного командного языка. Поддерживает множество форматов исполняемых файлов для большого числа процессоров и операционных систем [4]. Среда рассчитана на подключение отдельных модулей, расширяющих функциональность среды. Последние версии (IDA SDK) дают возможность пользователям самостоятельно разрабатывать модули. Разрабатываемые модули могут быть отнесены к одному из трех классов:
  - процессорные модули – позволяют разбирать программы для новых процессоров;
  - загрузчики – необходимы для поддержки новых форматов входных файлов;
  - плагины – модули, расширяющие функциональность IDA (добавление новых команд, улучшение анализа посредством установления обработчиков событий [5])

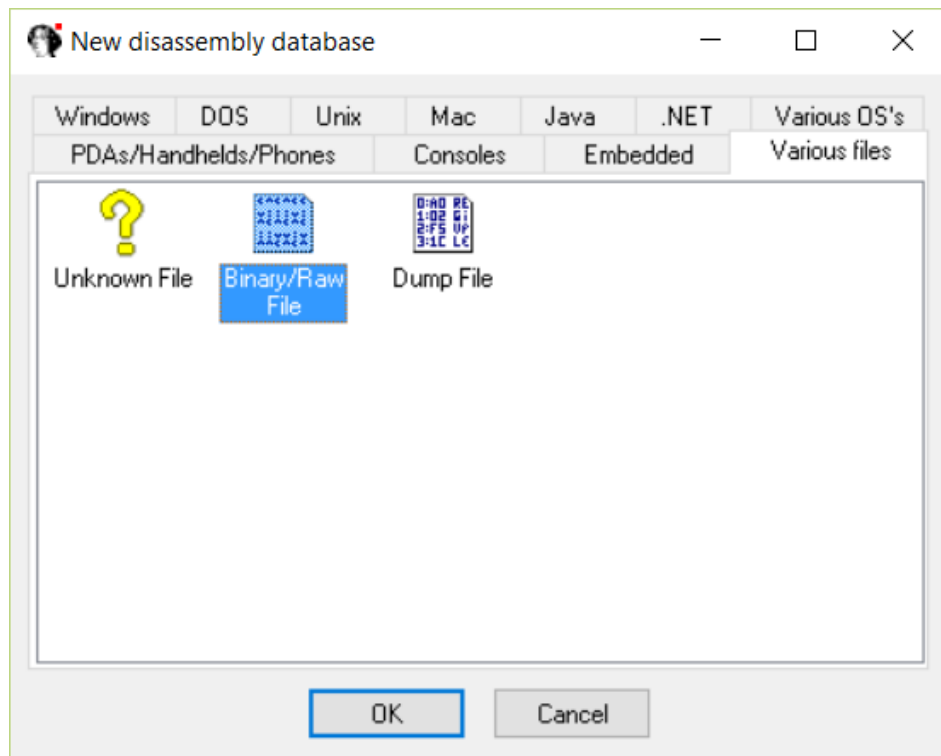
На данный момент IDA является лидером среди дизассемблеров на мировом рынке. Позволяет в автоматическом и полуавтоматическом режиме проводить дизассемблирование. Продвинутый уровень интерактивности позволяет изменять названия функций и переменных, сохраняет позицию курсора при перезагрузке среды, записывает изменения в базу данных (с возможным экспортом для другой ПК с IDA), позволяет выполнять переходы по адресам двойным кликом мыши. Кроме того, следует отметить, что существуют как платные, так и бесплатные версии данной среды. Это позволяет ее использовать всем желающим. Создателем среды является Ильфак Гильфанов (с группой друзей-программистов), разработчик с многолетним опытом и специалист по компьютерной безопасности [5].

Из существующих дизассемблеров, Interactive DisAssembler наиболее подходит для решения поставленной задачи. IDA не только обладает рядом функциональных преимуществ перед конкурентами, но и поддерживает дизассемблирование программ, написанных для искомого процессора TMS320C54.

### 3.1.2. Процесс дизассемблирования прошивки

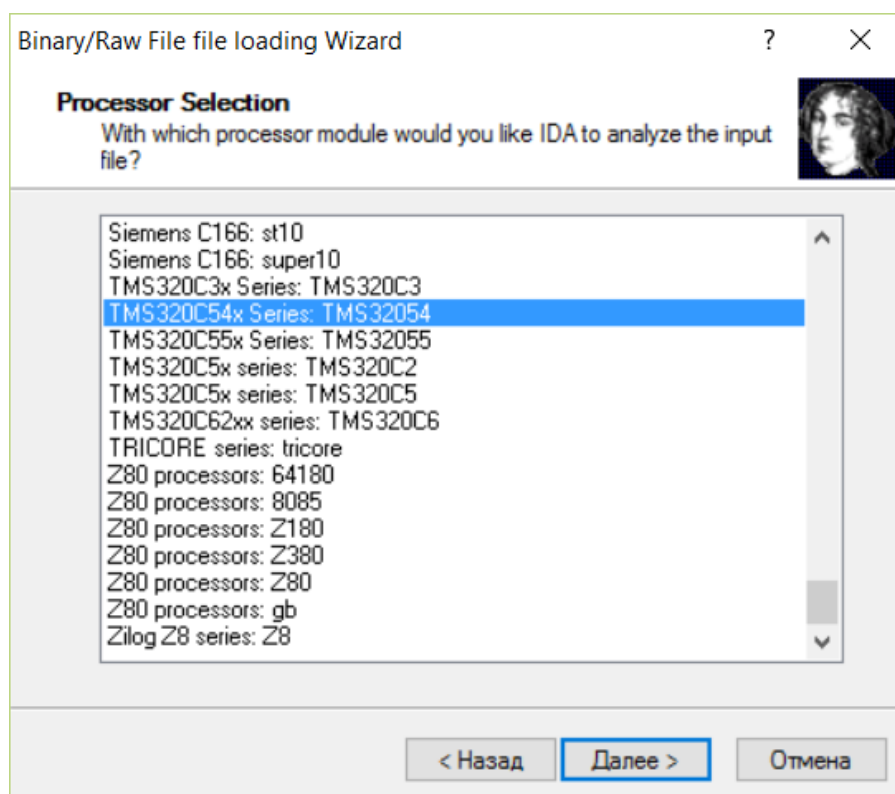
Для дизассемблирования прошивки необходимо выполнить следующие действия:

1. Запустить среду IDA и выполнить следующие команды **File → New → Binary File**. Откроется окно выбора типа входного файл, как на Рис. 4. После подтверждения типа файла, откроется окно в котором необходимо выбрать сам файл прошивки.



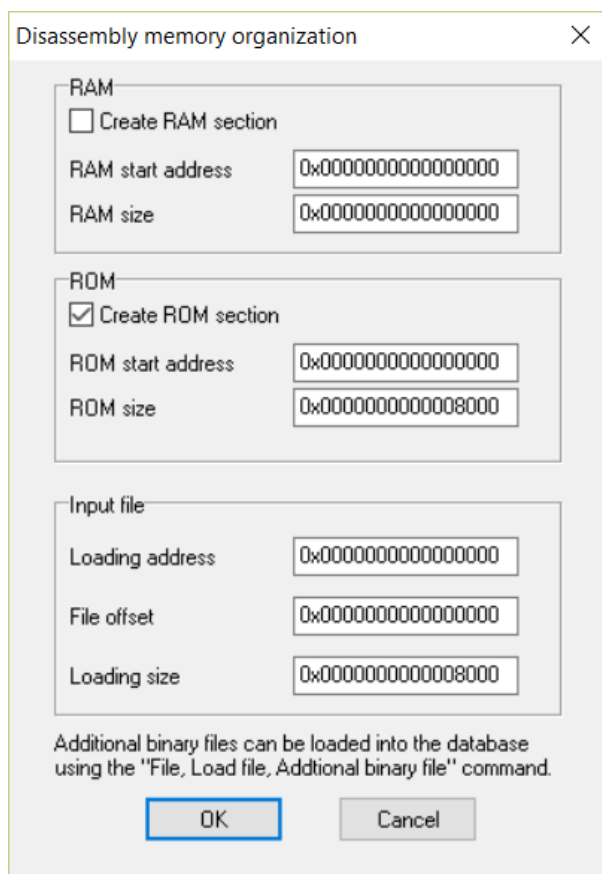
**Рис. 4. Выбор типа входного файла**

2. Затем откроется окно выбора типа процессора. Необходимо выбрать **TMS320C54** (см. Рис. 5).



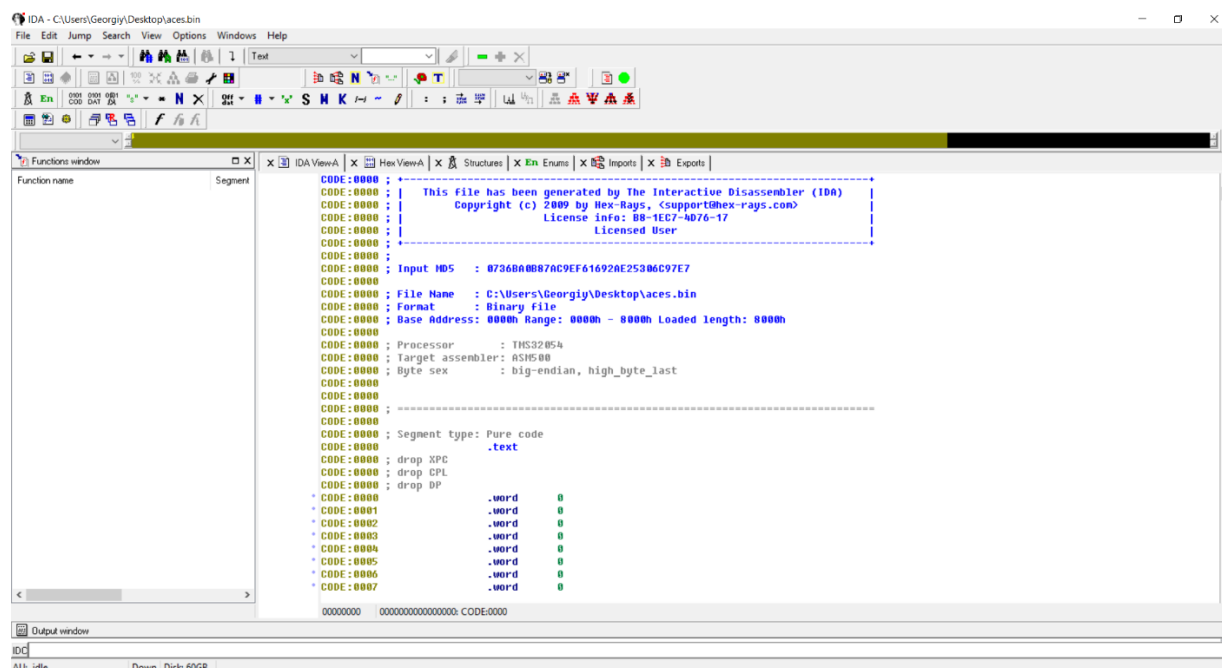
**Рис. 5. Выбор нужного процессора**

3. После выбора процессора **IDA** предложит указать участок памяти, куда будет загружена прошивка (см. Рис. 6). Так как задача - получить искомый код прошивки – никаких изменений вносить не нужно, необходимо оставить все по умолчанию.



**Рис. 6. Настройки адресов загрузки**

4. Сразу после завершения настроек откроется окно IDA View-A (см. Рис. 7), в котором и будет отображаться дизассемблированный код.



**Рис. 7. Основное окно разработки IDA View**

5. В открывшемся окне предоставлено содержимое бинарного файла. Необходимо установить курсор на нулевой адрес. По нажатию на кнопку "С" - небольшие участки кода начнут преобразовываться в ассемблерные команды, которые можно будет анализировать (см. Рис. 8). Перемещаясь по адресному пространству, необходимо выполнять преобразование бинарного кода в ассемблерные команды. IDA самостоятельно подписывает начало и конец функций, присваивая названия по адресу, например *sub\_2010*. Среда так же позволяет изменять названия функций и добавлять комментарии. Если пользователь проанализировал и понял некоторую функцию, он может изменить ее название на более подходящее или добавить необходимый комментарий двумя кликами ПКМ. Проводить статический анализ кода, в котором часть функций подписана и прокомментирована – в разы удобнее и быстрее, нежели чистый ассемблер файл.

```

CODE:200D      .word 140Ah
CODE:200E      .word 0F073h ; sE
CODE:200F      .word 200Eh
CODE:2010
CODE:2010 ; ===== S U B R O U T I N E =====
CODE:2010 ; Attributes: noreturn
CODE:2010
CODE:2010 sub_2010: ; CODE XREF: CODE:5E53↓p
CODE:2010 ; CODE:5E56↓p
CODE:2010      adds    *(byte_140A), A
CODE:2012      ld      #1400h, 0, B
CODE:2014      and     #0FFFFh, B
CODE:2016      sub     A, B
CODE:2017      rcd     bgeq
CODE:2018      stl     A, *(byte_140A)
CODE:201A

```

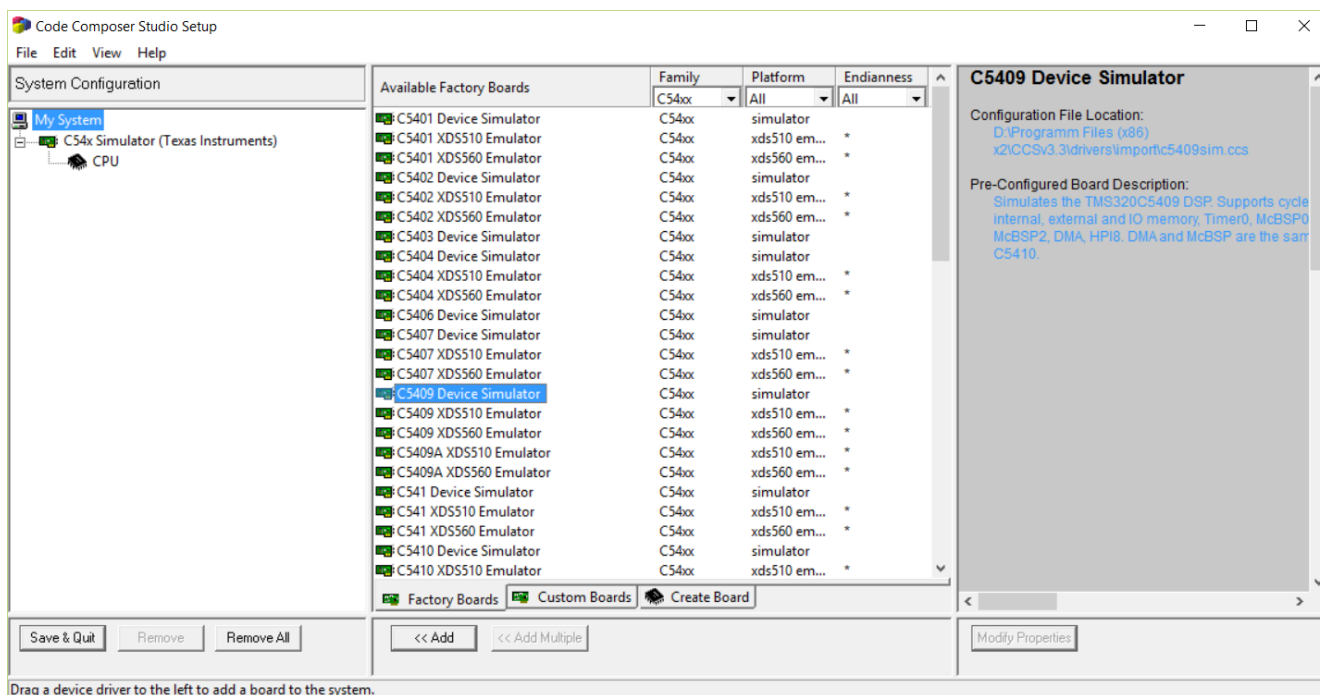
Рис. 8. Дизассемблированный участок кода

### 3.1.3. Выбор симулятора микропроцессора

Для того, чтобы найти кодек среди всей прошивки, одного статичного анализа недостаточно. Для удобного анализа кода потребуется симулятор микропроцессора. Симулятор позволит анализировать изменения регистров, аккумуляторов и памяти, что упростит первостепенную задачу – задачу обнаружения кодека. В связи с тем, что микропроцессор TMS320C54 является разработкой компании Texas Instrument [6], в качестве симулятора будет использоваться Code Composer Studio [7] – симулятор, разработанный той же фирмой для тестирования своих собственных процессоров. Данная среда зарекомендовала себя на рынке достойным образом, поэтому выбор очевиден. Существуют как платные версии, так и бесплатные с ограниченным набором микропроцессоров. TMS320C54 попадает как в платную версию программы, так и в бесплатную.

### 3.1.4. Настройка среды симулятора

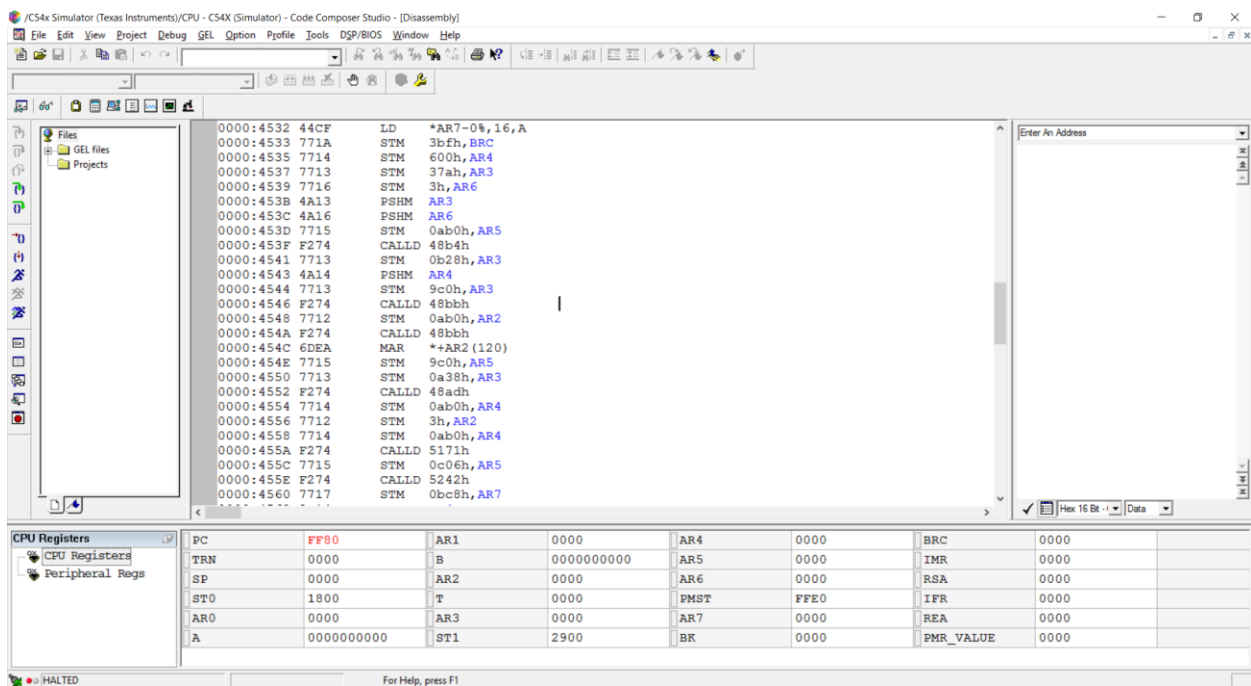
Среда состоит из двух модулей: *CCStudio(CCS)* и *CCStudio Setup(CCSS)*. Перед использованием симулятора необходимо установить тип симулируемого процессора в CCSS (см. Рис. 9)



**Рис. 9. Настройка CCS**

Необходимо найти процессор *TMS320C54* и добавить его к симулируемым (<<Add), после чего перезапустить CCS (Save&Quit). Среда перезапуститься и откроется основное окно CCS. Отобразить нужные окна, можно установив соответствующие модули во вкладке *View*. Потребуется следующие три окна: окно памяти, окно разборки и окно основных регистров (см. Рис. 10).





**Рис. 10. Необходимые окна для разработки в CCS**

После настройки среды необходимо загрузить саму прошивку и данные в симулятор. Дизассемблированная прошивка подается на вход CCS в виде dat-файла, содержащем необходимые заголовки, ассемблерные команды и данные в 16-ом коде. IDA способна сгенерировать файл содержащий все команды в 16-ом коде, однако ни заголовки, ни формат данных не подходит для CCS. Для корректной работы на вход CCS необходимо подать файл с 16-битными словами формата 1 (см. Рис. 11), с следующими заголовками:

- 1651 — корректная сигнатура для правильного распознавания формата данных и считывания CCS;
- 1 — формат подаваемых чисел. 1 — это hex формат. Команды можно подавать и в другом формате: 2 - int, 3-long, 4-float;
- 0 - начальный адрес (загрузка с нулевого адреса);
- 0 - текущая страница (нулевой адрес, следовательно, нулевая страница);
- FFFF - конечный адрес (крайний адрес в 16-битной системе) [8].

1	1651 1 0 0 ffff
2	0x0000
3	0x0000
4	0x0000
5	0x0000
6	0x0000
7	0x0000
8	0x0000
9	0x0000
10	0x0000
11	0x0000
12	0x1100
13	0x0000

**Рис. 11. Формат dat-файла для CCS**

IDA способна сохранить результат дизассемблирования в 16-битном формате (окно **hex view-A** → выделить код → ПКМ → **copy to file**) типа 2 (см. Рис. 12).

```

2185  7313 0062 F010 0003 8061 ED01 FA43 2191 E589 4592 4761 CB89 8793 F000 0002 8061
2195  8816 F000 0001 8814 7664 01FF 11E4 6524 8163 F77F 8910 7715 0011 7212 0060 7213
21A5  0062 721A 0061 F420 F272 21B4 7711 0040 6F85 0D59 1964 8914 6DB1 3093 28E4 6000
21B5  F00F 0001 8292 4910 6E8E 21A4 0163 8910 FC00 F000 0014 8060 F274 2318 F062 0008
21C5  FC00 F110 0018 FA4E 21D9 F020 6666 F300 0009 FA4F 21D9 F020 4CCC F300 000F 890E
21D5  F020 2666 F067 028F FC00 F6B8 F310 0001 891A F310 0008 890E F300 000F 44F8 000B
21E5  F48C F47F F57E F300 6320 8913 F030 0003 F562 890E 7714 0008 F004 FFFF F272 2201
21F5  1593 F764 F000 0001 962D FA20 2201 F180 8392 1193 F010 0004 F764 F7B8 FC00 F6B8
2205  F6B9 F310 0009 F761 F300 652F 8913 771A 0007 F272 2215 1093 3C83 F130 000F 8192
2215  F47C FE00 F7B8 F7B9 7311 0065 8064 8163 7312 0062 7313 0061 8060 F274 2318 F062
2225  0004 8066 3066 2061 F468 F00F 0001 8267 1064 F010 0001 881A 8064 7712 0068 F272
2235  2247 F420 F7B6 0067 F130 03FF 890E F576 0162 8913 2163 6F69 0D65 4563 4369 8368
2245  A598 B394 8391 F6B6 1065 8812 8813 6F66 0C5E 880E F420 4764 2892 F00F 0001 4409
2255  E732 ED00 4593 F520 4764 C798 FC00 7312 0066 7313 0068 7314 0069 E808 F074 2000
2265  806A F7B6 7212 0068 E808 F274 2183 7213 006A F420 8067 6F68 0D00 8914 6F69 0D00

```

**Рис. 12. Формат сохранения результата от IDA**

Конвертация формата 2 (IDA) в формат 1 (CCS) выполняется при помощи небольшой программы, написанной на Си (приложение 1). Результатом работы программы станет dat-файл с необходимыми заголовками и нужным форматом данных, который будет подаваться на вход CCS.

### 3.2. Анализ полученного кода

Перед вторым этапом разработки был поставлен целый ряд непростых задач:

- Найти функции кодека среди прошивки;

- Подтвердить работоспособность кодека, полученного после дизассемблирования;
- Провести анализ структуры кодека и его составляющих.

### 3.2.1. Поиск функции кодека

Для упрощения задачи поиска функций кодека в прошивке, потребовалось еще несколько вариантов прошивки этого же процессора. Каждая из прошивок отличается внутренним состоянием области памяти. Удобно было взять:

- Состояние памяти, до и после кодирования;
- Состояние памяти, сразу после принятия данных и после декодирования.

Поиск некоторого логически связанного участка прошивки рассмотрим на примере кодера. Анализируя прошивку с данными до кодирования представилось возможным найти адреса, в которые записываются входные данные. Функции, которые обращаются к этим адресам были выделены в группу 1. Таких функций оказалось не много, так как прошивка занимает примерно четверть всей внутренней памяти микропроцессора, в следствии чего свободной памяти на кристалле в избытке, и функции редко обращаются к одним и тем же адресам без логической связи друг с другом. Затем была проанализирована прошивка, содержащая закодированные данные, которую получили, взяв слепок памяти после кодирования. Функции, которые «что-то» пишут в область памяти, где расположились закодированные биты, были выделены в группу 2. Сопоставив функции групп 1 и 2, обнаружились функции, которые одновременно обращаются и к первым и ко-вторым адресам – это и есть основные функции кодера, точнее некоторая их часть (они были выделены в новую группу 3). Следующим шагом стало точное определение границ кодера. Необходимо было определить, какие близлежащие функции являются кодером, а какие нет. Выполняя и анализируя код функций группы 3 и прилежащих функций с помощью симулятора CCS, удалось с точностью до команды определить границы кодера. CCS не обладает интерактивностью, поэтому работу CCS удобно было комбинировать с IDA, оставляя пометки и

комментарии в среде дизассемблера. Поиск декодера выполнялся подобным образом, на основании прошивок до и после декодирования.

### **3.2.2. Подтвердить работоспособность кода**

После того, как границы кодера и декодера были определены, потребовалось подтвердить работоспособность кода, полученного после дизассемблирования. Тестировать отдельные участки кода (например, кодер или декодер) оказалось намного проще, нежели всю прошивку целиком. Именно поэтому этап проверки кода идет только после обнаружения границ кода.

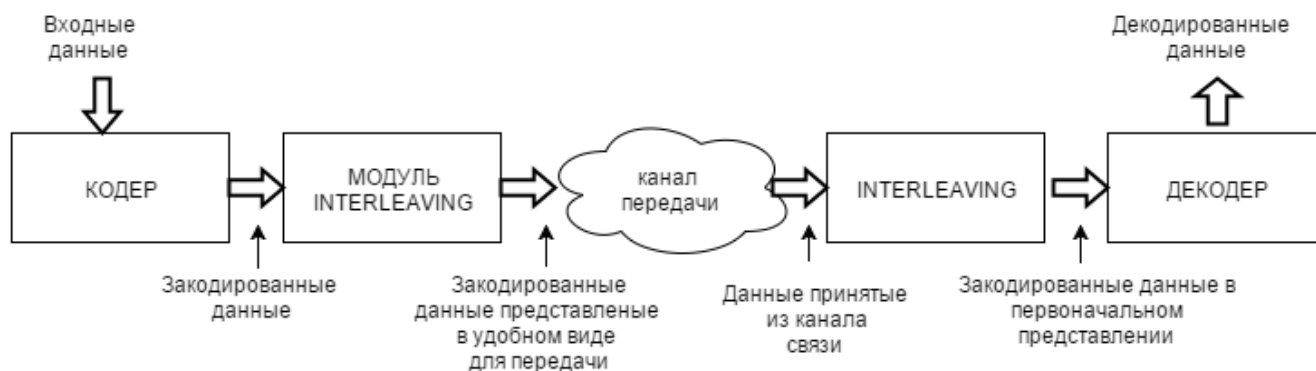
В связи с тем, что IDA поддерживает TMS320C54, вероятность возникновения ошибки при дизассемблировании сводиться к нулю. Однако писать собственные модули, не проверив работоспособность ранее полученного кода не разумно - «Здоровое недоверие — хорошая основа для совместной работы» [9]. Для того, чтобы удостовериться в работоспособности кода, в симулятор (CCS) последовательно загружаются два dat-файла: первый - полученный код после дизассемблирования с помощью IDA, и второй - данные которые необходимо закодировать. Средствами симулятора выполняется кодирование/декодирование. Результат кодирования/декодирования должен совпадать с эталонной моделью (прошивки устройства в соответствующих состояниях) не только по содержанию (данные бит в бит), но и по адресам расположения данных. Т.к. задача стоит подтвердить корректность кода после дизассемблирования, нескольких тестов с успешным исходом хватит (при условии, что отсутствуют тест с неуспешным исходом). Ошибки, возникающие в коде в процессе обратной разработки чаще всего, несут глобальный характер, как следствие уже первые тесты показывают некорректность кода. В таком случае необходимо проверить корректность искомой прошивки и правильность настроек. Большинство ошибок допускаются именно на этих этапах, конкретного универсального решения таких проблем нет. Возможно стоит

повторить весь процесс дизассемблирования с начала или же попробовать отладить код самому, что без понимания практически невозможно.

В случае, если проблем не возникло, или они были успешно решены, можно переходить к следующему этапу.

### 3.2.3. Анализ структуры кодека

После завершения дизассемблирования и тестирования работоспособности кода, был проведен полный анализ функций и алгоритмов кодирования и декодирования. В результате было выяснено, что кодек состоит из 12 функций и подразделяется на нескольких основных модулей: свёрточный кодер, модуль прямого чередования (interleaving), модуль обратного чередования (deinterleaving), декодер- Витерби.



**Рис. 13. Структурная схема работы кодека**

Структура кодека приведена на Рис. 13. Ниже рассмотрим смысловую часть каждого из модулей.

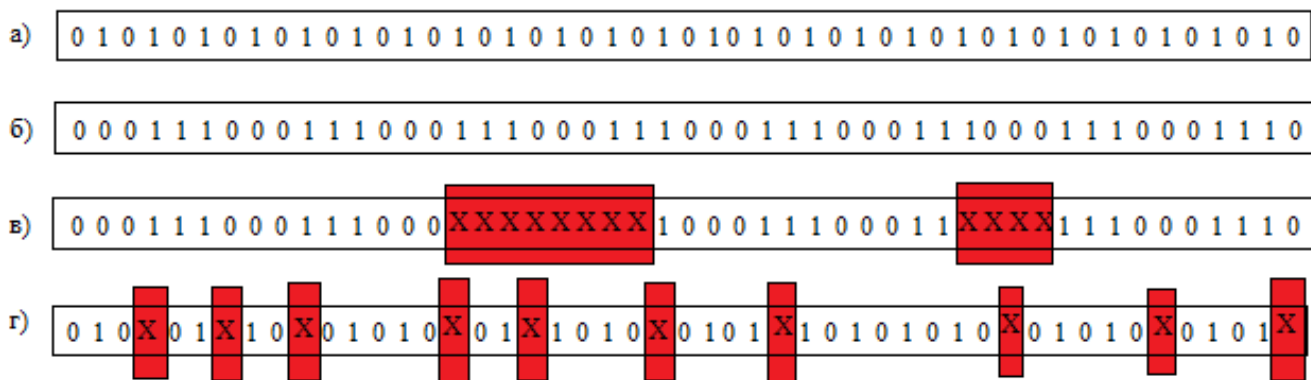
#### 3.2.3.1. Кодер

Кодер выполняет кодирование указанного участка памяти в одном из двух режимов: кодирование двумя битами или четырьмя. По завершению кодирования каждый из кодируемых бит будет представлен двумя или четырьмя битами соответственно, то есть выполняется сверточное кодирование. В основе данного модуля лежит 4 полинома. Для режима кодирования двумя битами используется только два полинома, для

кодирования четырьмя битами – все четыре. Полиномы зависят от предыдущих шести бит - это означает, что кодер с памятью, то есть это конечный автомат. Количество возможных комбинаций ограничено 64 возможными, значит в конечном автомате 64 состояния.

### 3.2.3.2. Модули чередования

После того, как данные успешно закодированы, они подвергаются чередованию. Этот процесс необходим для снижения вероятности неудачного декодирования. Во время передачи данных, вследствие воздействия шумов происходят битовые ошибки, часть данных теряется или принимается неверно. Декодер способен восстановить часть потерянных данных за счет свойств кода, однако качество декодирования и распознавания ошибок, напрямую зависит от количества ошибочных бит, следующих подряд друг за другом (пачки ошибок). В канале же при возникновении зашумления – портится целая последовательность бит, длиной от нескольких десятков до нескольких сотен бит.



**Рис. 14. а) Данные после кодирования; б) Данные после чередования;  
в) Данные после передачи по каналу связи; г) Данные после ликвидации  
чередования.**

Для того, чтобы минимизировать длину последовательностей, ошибочных бит, применяются прямая и обратная операция чередования. Модуль чередования перемешивает биты в соответствии со стандартным алгоритмом. Затем данные поступают в канал передачи, откуда считываются

другим устройством. На принятом устройстве закодированные данные поступают в модуль обратного чередования, который восстанавливает исходную последовательность бит, которые в свою очередь уже могут быть декодированы. На Рис. 14 приведен простой пример работы данных модулей.

Алгоритмы чередования не представляют большого интереса, поэтому подробно не представлены. Можно только сказать, что основной задачей чередования является максимальный разброс соседних бит посылке.

### **3.2.3.3. Декодер**

На последнем этапе закодированные данные поступают в декодер, где согласно алгоритму Витерби производится декодирование. Стоит отметить, что данная реализация декодера-Витерби является реализацией с мягкими весами, т.е. на вход декодера поступают биты с весами, которые присваиваются сторонними компонентами. В данном случае вес – это вероятностная мера, утверждающая с некоторой точностью, что принятый бит – это ноль или единица.

Тело алгоритма состоит из двух основных частей: прямой и обратной проходки. При прямой проходке выстраивается решетка Витерби – записываются метрики возможных путей декодирования. Решетка шириной 64 состояния, глубина декодирования составляет 6. Это означает, что на каждом шаге прямой проходки алгоритм отслеживает лучшие 64 пути. При обратной проходке, на основе полученных метрик, определяется путь с минимальной ошибкой. Выполняя проходку по решетке в обратном направлении, закодированное сообщение восстанавливается.

## **3.3. Разработка модулей кодека на языке Си**

На основе логических связей функции были разбиты на 4 группы. Каждая из групп послужила образованию отдельного модуля: сверточного кодера, двух модулей чередования и модуля декодера. Перед каждым модулем поставлена отдельная задача, которую он должен выполнять максимально эффективно.

### 3.3.1. Кодер

С целью повышения эффективности, все возможные результаты работы полиномов рассчитаны заранее и сохранены в двумерном массиве 4x8, в виде 16-битных чисел. Каждый столбец соответствует конкретному полиному, а номер строки и бита – результату работы полинома. В зависимости от того какие биты были приняты на вход кодера ранее, определяется необходимая строка и бит. Для сверточного кодера был реализован следующий алгоритм (см. Рис. 15):



**Рис. 15. Структура работы кодера**

Тело кодера состоит из двух вложенных циклов. В теле первого цикла происходит определение номера строки и бита, в теле второго - копирование необходимого бита для каждого из полиномов. Количество используемых полиномов зависит от режима кодирования.



### 3.3.2. Модули Чередования

На вход в качестве параметров модули принимают указатель на память и длину сообщения. Данные так же записываются в память на некотором смещении.

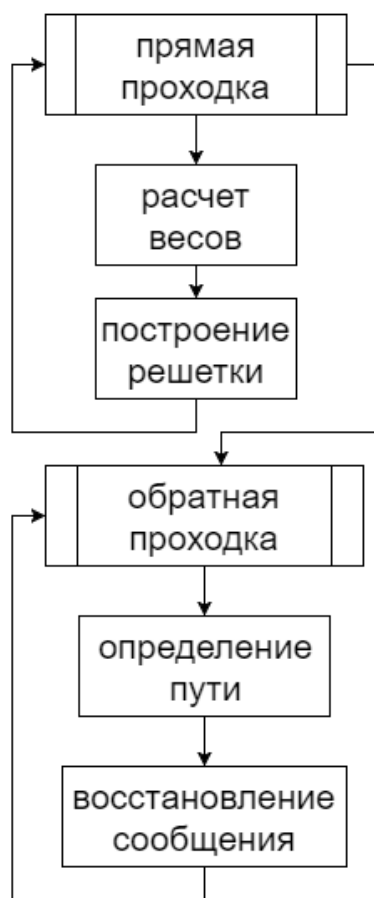
Алгоритм чередования состоит из двух частей (см. Рис. 16): диагонального и вертикального смешивания. Данные при этом представляются в виде условной матрицы  $N \times 120$ , где каждая строка – это отдельный участок сообщения.



**Рис. 16. Структура работы кодера**

При диагональном чередовании выполняется считывание по диагоналям, в результате данные представляются в виде новой матрицы. Затем, при вертикальном смешивании выполняется транспонирование матрицы, полученной после диагонального смешивания. Результатом работы алгоритма становится перемешивание сообщения с двумя соседними. В случае же, если это было первое сообщение или последнее, в передаваемом потоке данных, то оно смешивается с нулевым пакетом (сообщение с нулями), который будет отброшен после передачи.

### 3.3.3. Декодер



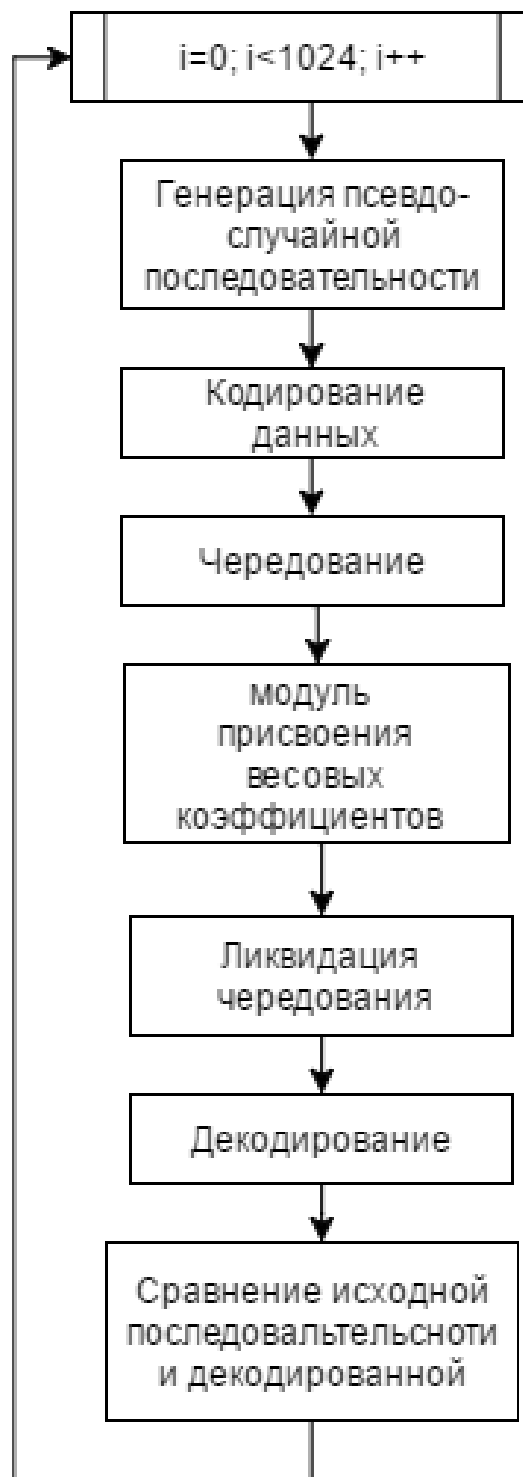
**Рис. 17. Алгоритм декодирования**

Модуль декодер состоит из 3 функций: 1 основной, которая включает в себя прямую и обратную проходку (см. Рис. 17), и 2-ух вспомогательных функций, которые рассчитывают все варианты ошибок на закодированном участке и соответствующие веса, необходимые для постройки решетки Витерби. Для решетки создается динамический массив TRN, размерность которого зависит от количества декодируемых бит. Решение о переходе состояний решетки принимается на основе рассчитанных весов закодированного участка бит. Метрики возможных путей декодирования при этом сохраняются в массиве TRN. При обратной проходке на основе метрик из массива TRN определяется маршрут с наименьшей накопленной ошибкой. Выполняя проходку по решетки на основе выбранного пути, данные восстанавливаются (с конца в начало).

Для простоты использования, кодек был оформлен в качестве библиотеки (Приложение 2). При этом каждый модуль представлен в виде отдельной функции или группы функций. Модуль-образующие функции принимают на вход: указатели на участки памяти (для чтения и записи), длину участка памяти, который необходимо подвергнуть обработке, и режим кодирования, необходимый только кодеру и декодеру. При разработке использовался только язык Си. Были подключены две сторонние библиотеки: `<math.h>` – для команды `pow` (возведение в степень) и `<stdlib.h>` - для выделения памяти под массив TRN в декодере.

### **3.4. Тестирование реализации на языке Си**

По завершению разработки было проведено тестирование. Во-первых, необходимо было подтвердить факт работоспособности разработанного кодека, во-вторых подтвердить совместимость разработанного кодека и устаревшего устройства. Для того, чтобы выполнить поставленные задачи потребовалось написать отдельную программу (приложение 3), которая запускает модули кодека в необходимом порядке, передавая на кодирование различные данные, сгенерированные псевдослучайным образом. Алгоритм разработанной программы показан на Рис. 18.



**Рис. 18. Алгоритм тестирования кодека**

Работа программы происходит следующим образом:

В цикле генерируется случайная последовательность нулей и единиц. Затем сгенерированная последовательность передается на вход кодера, где она подвергается кодированию и чередованию. После кодирования данные готовы

для отправки в канал. Модуль весовых коэффициентов присваивает однозначные веса передаваемым битам (Приложение 4). После получения данных из «канала», происходит ликвидация чередования и декодирование данных. По завершению декодирования полученная и искомая последовательности сравниваются. В случае если они не совпали, номер теста и ключ последовательности выводятся на экран. Анализируя провалившиеся тесты с помощью отладчика и симулятора выполняется поиск и исправление допущенных ошибок в коде.

Результатом первого этапа тестирования стало обнаружение одной ошибки в реализации на языке Си. После ее исправления, тестирование было проведено повторно (с добавлением неудавшихся тестов) и ошибок не выявило.

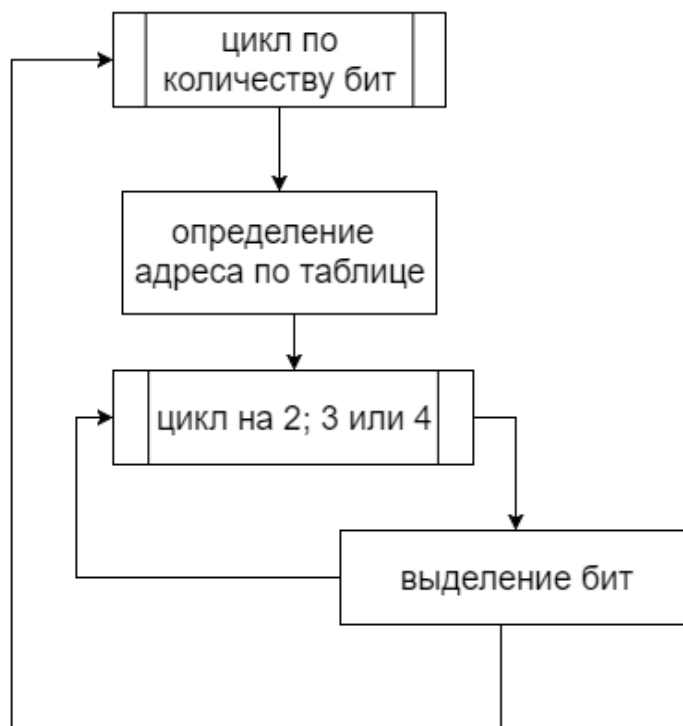
На втором этапе тестирования необходимо было подтвердить возможность взаимодействия старого и нового устройства. В виду того, симулятор невозможно включить в цикл тестирования (т.к. CCS является GUI приложением), было проведено выборочное тестирование 50 последовательностей. Данные кодировались кодером с прошивки, средствами симулятора CCS, и кодером написанном на языке Си, средствами компилятора Си, после чего результаты сравнивались. Различий обнаружено не было, из чего можно сделать вывод, взаимодействие между старым и новым устройством возможно.

### **3.5. Модернизация кодека**

Очевидный недостаток кодека, изъятых из прошивки – это отсутствие промежуточного режима кодирования кодирование тремя битами. Наличие такого режима позволит сократить временные затраты на кодирование, передачу и декодирование в целом ряда случаев.

### 3.5.1. Кодер

Внедрение промежуточного режима должно минимально изменять структуру модулей, поэтому использовались первые три полинома из существующих. При этом, вся модернизация кодера заключается в изменении кол-ва возможных вложенных циклов (см. Рис. 19).



**Рис. 19. Алгоритм кодера с промежуточным режимом.**

### 3.5.2. Модули Чередования

Модули прямого и обратного чередования изменениям не подвергаются, так как они не зависят от режима кодирования.

### 3.5.3. Декодер

Для модернизации декодера необходимо было решить следующие две задачи: во-первых, необходимо было разработать вспомогательную функцию для нового режима, которая должна была бы, аналогично двум другим, рассчитывать веса возможных ошибок для построения решетки Витерби. А во-

вторых, рассчитать переходы по конечному автомату для используемых полиномов.

Для решения первой задачи потребовалось проанализировать алгоритмы двух существующих функций. Рассмотрим алгоритм работы этих функций на примере режима 2. На вход функции поступает последовательность из двух бит, то есть один закодированный. Каждый принятый бит представляется 8 битами: один – под значение, семь – под весовой коэффициент, присвоенный после принятия бита из канала передачи. Функция разделяет значения и веса. Конкатенированные значения обоих бит образуют *дугу перехода* (со значением от 0 до 3). На основе выделенных весов рассчитываются все варианты ошибок. В таблице 1 представлен пример расчета для режима 2.

**Таблица 1. Пример расчета веса ошибок**

1 бит	2 бит	Итоговый вес:
-	-	Вес 1 бита + Вес 2 бита
-	+	Вес 1 бита - Вес 2 бита
+	-	Вес 2 бита - Вес 1 бита
+	+	- Вес 1 бита - Вес 2 бита

Знак «+» означает, что бит принят верно, знак «-» - бит принят с ошибкой. После того, как все варианты ошибок были рассчитаны, в решетке выполняется полный перебор возможных путей. Полученная ранее *дуга перехода* сравнивается с возможными переходами из каждого текущего состояния решетки. Если один (или два) бита не совпадают, то к пути добавляется рассчитанное значение ошибки. Если же оба бита совпали – то от веса всего пути отнимается вес правильных бит. Такая организация работы цикла, построения решетки, позволяет сократить внутренние вычисления, так как вместо того, чтобы считать веса для каждого конкретного случая, вычисления проводятся вне цикла один раз.

Для решения второй задачи пришлось так же проанализировать последовательности переходов по конечному автомату для режима 2 и 4. Конечные автоматы для всех трех случаев одинаковые, однако числа, которыми закодированы дуги – отличаются. Из анализа чисел стало понятно, что каждый бит числа – это результат работы одного из полиномов. Это позволило рассчитать веса дуг переходов для используемых полиномов.

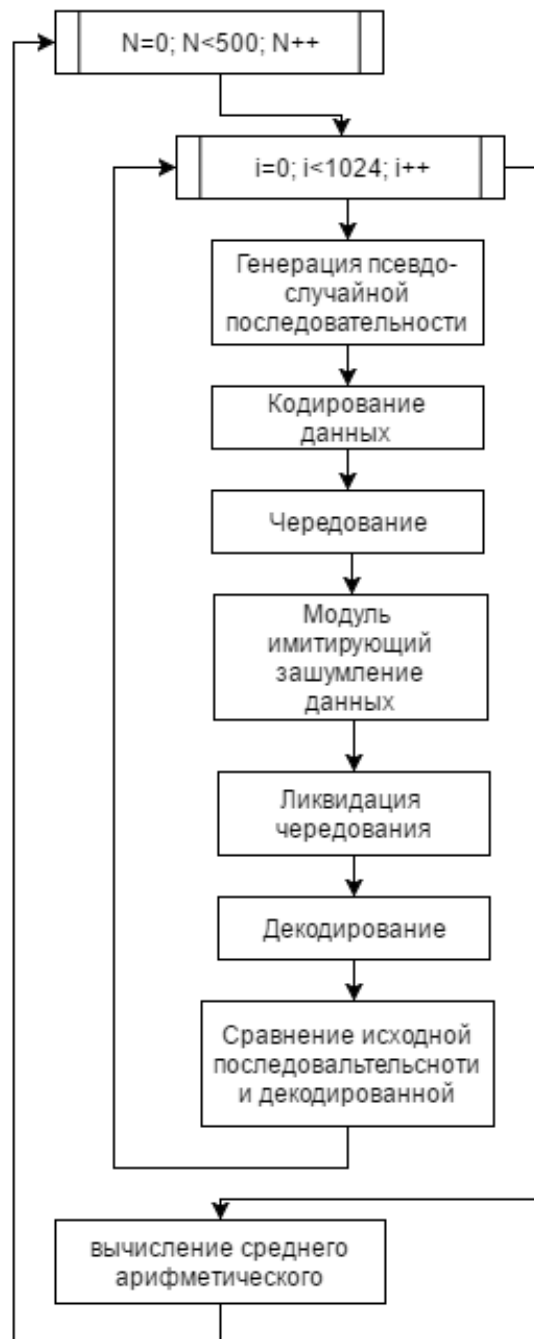
Структура алгоритма при этом никак не изменилась.

По завершению разработки было проведено тестирование как уже существующих режимов, на случай если возникли случайные ошибки при разработке, так и нового режима. В результате было выявлено две ошибки внутри разработанного режима. После их исправления – тесты подтвердили работоспособность нового режима. Для тестирования использовалась программа из пункта 3.4.

### **3.6. Тестирование декодирующих способностей всех трех режимов**

На следующем этапе было необходимо проверить и подтвердить восстанавливающие способности разработанного декодера и определить граничные значения восстанавливающих способностей. Для этого программа из предыдущего пункта тестирования (приложение 3) была модифицирована, а модуль присваивающий однозначные весовые коэффициенты (приложение 4) был заменен на модуль имитирующий шум канала (приложение 5).





**Рис. 20. Алгоритм тестирования декодирующих способностей**

Алгоритм программы представлен на Рис. 20. Внутренний цикл генерирует случайную последовательность, после она кодируется. В закодированное сообщение вносятся  $N$  ошибок (определяется значением из внешнего цикла), затем декодируется. На основе информации о количестве допущенных ошибок на 1024 опыта высчитывается среднее арифметическое с округлением в большую сторону. Результат вычисления записывается в файл. На основе готового файла можно построить кривую, показывающую

зависимость восстанавливающих способностей от количества допущенных ошибок.

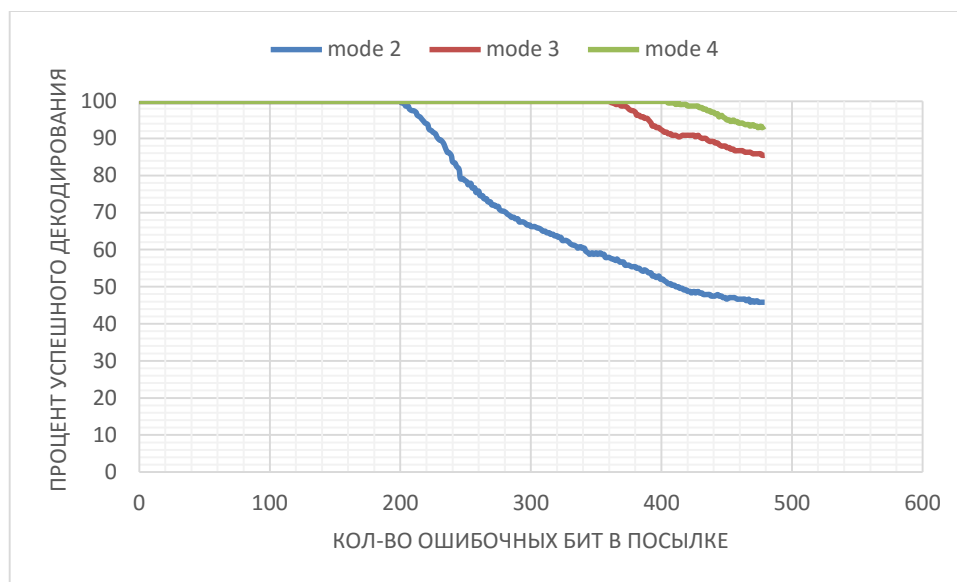
Модуль имитирующий шум имеет три входных параметра, задаваемых перед запуском модуля: количество зашумлённых участков в передаваемой посылке, длина первого зашумленного участка (если участков  $n$ , каждый следующий будет на  $(100/n)$  % короче первого) и расстояние между зашумленными участками (в случае, если данный параметр не указан – зашумленные участки создаются на максимальном расстоянии друг от друга, имитируя наихудший вариант. Это объясняется работой модуля чередования, т.к. он распределяет данные равномерно на максимальном расстоянии друг от друга). Во избежание получения однотипных результатов, модуль содержит ряд случайно сгенерированных внутренних параметров (например, расстояние, от начала посылки, на котором генерируется первое зашумление). Биты, которые необходимо подвергнуть зашумлению – инвертируются и им присваивается псевдослучайный вес.

Для тестирования восстанавливающих способностей декодера были проведены эксперименты со настройками модуля-шума, приведенными в таблице №2.

**Таблица 2. Настройки модуля шума для трех групп опытов**

№	Кол-во зашумленных участков	Кол-во испорченных бит	Расстояние между ОП
1	1	1-492	-
2	2	1-360; 1-180;	0-120
3	3	1-240; 1-120; 1-60;	0-120

В результате проведенных опытов зависимость восстанавливающих способностей от количества зашумленных участков не выявилась. Результаты всех опытов оказались схожи, и зависят только от количества ошибочных бит. В связи с этим на Рис. 21 предоставлен результат только для первой группы опытов.



**Рис. 21. Результат тестирования декодирующих способностей.**

На Рис. 21 предоставлена визуальная характеристика декодирующих способностей различных режимов кодирования. Эксперимент был поставлен следующим образом: генерируется порядка тысячи случайных последовательностей по 240 бит каждая. Последовательности кодируются, длина закодированного сообщения при этом составляет 1440; 1680; 1920 (т.к. модуль чередования смешивает несколько сообщений сразу, или если это единственное сообщение – смешивает с нулевым пакетом, 960 нулевых бит). Такое увеличение обусловлено работой модуля чередования. После чередования в каждую из закодированных посылок вносится  $N$  ошибок и сообщения декодируются. По завершению декодирования все последовательности сравниваются с исходными. В случае, если исходная и декодированная последовательность не совпали – подсчитывает кол-во несовпадений. Затем подсчитывается среднее арифметическое всех ошибок (с

округлением в большую сторону). На основе среднеарифметического вычисляется % успешного декодирования.

Из результатов видно следующее:

Режим 2 способен исправить до 201 ошибочных бит из 1440 бит передаваемого сообщения ( $240 * 2 + 960$  бит нулевого пакета), что соответствует 13,9% исправления ошибок.

Режим 3, добавленный в ходе разработки, способен удачно декодировать сообщение при 360 ошибках из 1680 передаваемых бит ( $240*3+$  нулевой пакет), что соответствует 21,4%.

Режим 4, удачно декодирует сообщение при немного большем, чем 400 ошибок из 1920 передаваемых бит, то есть при 20,8%.

Как видно, режим 3 практически не уступает в декодирующих способностях режиму 4, но выигрывает по скорости в 1,5 раза (нулевые пакеты не учитывают, в связи их бесконечно малым размером при полноценной передаче). Очевидно, что разработанный режим эффективен. Он позволит сократить затраты на кодирование, передачу и декодирование на 30-35%, для случаев где количество ошибочных бит колеблется от 200 до 360 бит на передаваемое сообщение. Таким образом, внедренный режим практически полностью вытесняет режим 4, за исключением тех случаев, когда количество ошибок на передаваемое сообщение изменяется от 360 до 400 бит.

## **Заключение**

В ходе работы был выполнен целый ряд сложных задач:

- Было проведено дизассемблирование прошивки устройства, базирующегося на микропроцессоре TMS320C54.
- В полученном коде были определены функции, принадлежащие кодеку.
- Был проведен тщательный анализ искомого кодека.

- На основе полученных знаний, был разработан и внедрен промежуточный режим кодирования, показывавший свою эффективность уже при первых тестах.
- Кодек был реализован на языке программирования Си и оформлен в виде библиотеки.

Разработанная библиотека будет применяться для создания прошивок новых устройств, базирующихся как на основе TMS320C54, так и на основе любых других микропроцессоров.

Структура разработанного кодека позволяет, в случае необходимости, внедрить дополнительные режимы, отличающиеся полиномами кодирования или количеством используемых бит для кодирования.

## Список использованных источников

1. Техническое описание микропроцессора TMS320C54: [Электронный документ]. (<http://www.ti.com.cn/cn/lit/ug/spru172c/spru172c.pdf>). Проверено 19.05.2016.
2. Статья «Алгоритм Витерби» на сайте «nlpub.ru»: [Электронный документ]. ([https://nlpub.ru/Алгоритм\\_Витерби](https://nlpub.ru/Алгоритм_Витерби)). Проверено 19.05.2016.
3. Статья «Hiew» на сайте «Wikipedia.org»: [Электронный документ]. (<https://ru.wikipedia.org/wiki/Hiew>). Проверено 19.05.2016.
4. Статья «IDA Pro - интерактивный дизассемблер» на сайте «www.idasoft.ru»: [Электронный документ]. (<http://www.idasoft.ru/idapro/>). Проверено 19.05.2016.
5. Интервью с Ильфаком Гильфановым, разработчиком IDA: [Электронный документ]. (<http://fcenter.ru/online/softarticles/interview/6704>). Проверено 19.05.2016.
6. Официальный сайт фирмы «Texas Instrument»: [Электронный документ]. (<http://www.ti.com/>). Проверено 19.05.2016.
7. Code Composer Studio (CCS) Integrated Development Environment (IDE): [Электронный документ]. (<http://www.ti.com/tool/ccstudio>). Проверено 19.05.2016.
8. Описание сигнатуры dat-файла: [Электронный документ]. ([https://e2e.ti.com/support/development\\_tools/code\\_composer\\_studio/f/81/t/277303](https://e2e.ti.com/support/development_tools/code_composer_studio/f/81/t/277303)). Проверено 19.05.2016.
9. И.В.Сталин в беседе с Б. Куном; С. 190.

## Приложение 1

### Листинг 1. Файл XXXX.c

## **Заключительный лист работы**

Выпускная работа выполнена мною самостоятельно. Использованные в работе материалы и концепции из опубликованной научной литературы и других источников имеют ссылки на них.

Список использованных источников содержит \_\_\_\_\_ наименований.

Работа выполнена на \_\_\_\_\_ страницах.

Приложения к работе на \_\_\_\_\_ страницах.

Один экземпляр сдан в директорат.

Подпись \_\_\_\_\_ / \_\_\_\_\_ /

(фамилия, инициалы)

Дата « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_\_ г.