

АДАПТИВНЫЙ АЛГОРИТМ СЖАТИЯ ДАННЫХ В КОМПЬЮТЕРНЫХ СЕТЯХ

к.т.н. И.В. Рубан, М.Н. Колмыков, А.А. Резуненко
(представил д.ф.-м.н., проф. С.В. Смеляков)

В данной статье проводится анализ существующих протоколов сжатия информации, применяемых в аппаратуре приема-передачи данных в компьютерных сетях, и предлагается адаптивный подход к сжатию информационных потоков.

Постановка проблемы. В настоящее время для обеспечения высокой оперативности передачи данных в компьютерных сетях в аппаратуре приема-передачи данных широко используются протоколы сжатия информации, реализованные в сертифицированные протоколы сжатия.

Существенным недостатком данных протоколов является то, что они, по сути, не различают виды информации, один и тот же алгоритм осуществляет сжатие текста, графики, видео, звуковых данных, хотя для обработки каждого вида информации существуют индивидуальные алгоритмы сжатия, позволяющие добиваться высокой степени сжатия на основе учета статистических свойств информации.

Анализ литературы. Исходя из анализа литературы, наиболее распространенными протоколами сжатия являются MNP5, MNP7, и V.42bis [1, 3 – 6]. Алгоритмы протоколов MNP основаны на относительно простых методах сжатия, его эффективность в лучших случаях редко превышает 2. Протокол V.42bis основан на популярном методе сжатия LZW, применяемом в большинстве архиваторов, и в удачных случаях обеспечивает сжатие лишь до четырех раз.

Цель статьи. Предложить метод сжатия информационных потоков на основе компактного представления данных в зависимости от вида информации и маркирования пакетов данных.

Принцип действия основных протоколов сжатия данных. Протокол MNP5 реализует комбинацию адаптивного кодирования с применением кода Хаффмена и группового кодирования. При этом хорошо поддающиеся сжатию данные уменьшают свой исходный объем примерно

на 50 % и, следовательно, реальная скорость их передачи возрастает вдвое по сравнению с номинальной скоростью передачи данных модемом.

Работа его заключается в использовании на первом этапе процедуры сжатия метода группового кодирования для удаления из потока передаваемых данных слишком длинных последовательностей повторяющихся символов (рис. 1), и на втором этапе – адаптивного кодирования на основе метода Хаффмена, основанного на предположении, что некоторые символы будут встречаться в потоке данных чаще, чем другие [2]. Символы, которые встречаются чаще, кодируются с использованием небольшого числа битов. Реже встречающиеся символы передаются с использованием более длинных кодовых последовательностей.

A	B	B	C	C	C	D	D	D	D	E	E	E	E	E	Перед сжатием
A	B	B	C	C	C	0	D	D	D	1	E	E	E	2	После сжатия

Рис. 1. Групповое кодирования потока данных на первом этапе алгоритма

Особенностью метода является:

- на первом этапе – проверка проходящего потока данных: в процессе работы алгоритм определяет одинаковые символы, после чего начинает счет удалять из потока данных до 250 одинаковых следующих друг за другом символов;

- на втором этапе, когда формат передаваемых данных относительно хорошо известен и постоянен, кодовые битовые последовательности, или лексемы, определяются заранее и адаптивный алгоритм подстраивается под поток с последующим изменением своих лексем.

Основным недостатком протокола MNP5 является то, что алгоритм сжатия не отключается, и протокол всегда пытается кодировать поступающие данные. Это часто приводит к тому, что данные, не поддающиеся сжатию, за счет кодирования увеличиваются в размере, и эффективная скорость передачи падает, так как кодирование группы из трех символов, наоборот, расширяет поток данных.

По сравнению с MNP5 протокол MNP7 использует более эффективный алгоритм сжатия данных и позволяет достичь коэффициента сжатия 3:1.

Особенностью протокола MNP7 является использование улучшенной формы кодирования метода Хаффмена в сочетании с марковским алгоритмом прогнозирования для создания кодовых последовательностей минимально возможной длины.

Принцип действия алгоритма показан на рис. 2 и заключается в том, что марковский алгоритм может предсказывать следующий символ в последовательности, исходя из появившегося предыдущего символа. Для

каждого символа, или октета, формируется таблица из всех 256 возможных следующих за ним октетов, расположенных в соответствии с частотой их появления. Октет кодируется путем выбора столбца, соответствующего предыдущему октету (озаглавливаемому столбец), с последующим отысканием в этом столбце значения текущего октета. Строка, в которой находится текущий октет, определяет лексему точно так же, как в описанном выше случае кодирования с использованием кода Хаффмена. После того, как каждый октет будет закодирован, порядок следования записей (октетов) в выбранном столбце изменяется в соответствии с новыми относительными частотами появления октетов.

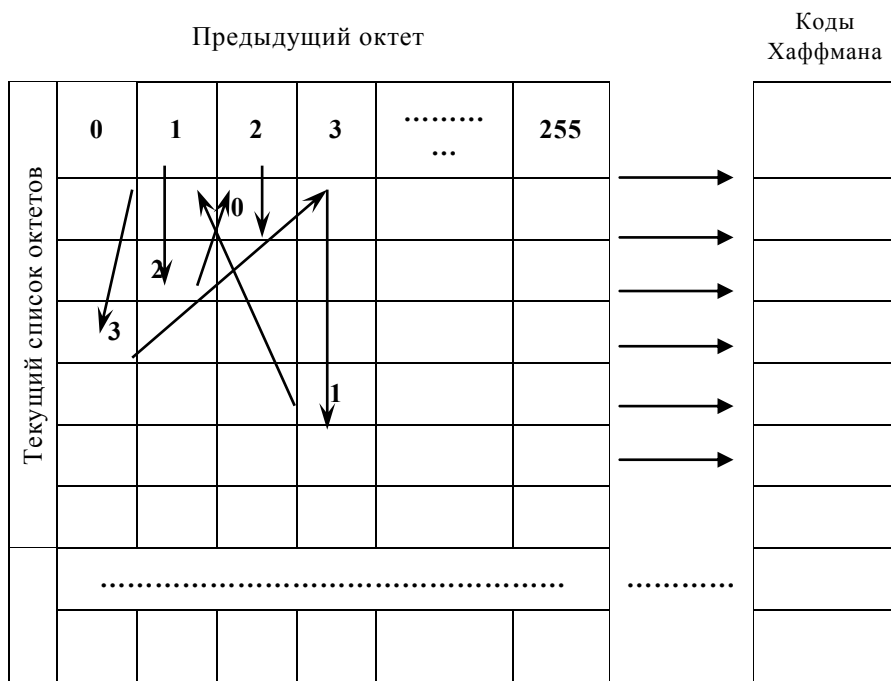


Рис. 2. Кодирование при помощи марковского алгоритма Хаффмена

В отличие от протоколов MNP другой самый распространенный протокол модемного сжатия V.42bis реализует модификацию метода словарей Якйба Зива и Абрахама Лемпеля (LZW).

Особенность метода сжатия данных LZW состоит в создании древовидного словаря последовательностей символов, в котором каждой последовательности соответствует единственное кодовое слово. Входящий поток данных последовательно, символ за символом, сравнивается с

имеющимися в словаре последовательностями. После того, как в словаре будет найдена кодируемая последовательность, идентичная входной, модем передает соответствующее ей кодовое слово. Алгоритм динамически создает и обновляет словарь символьных последовательностей.

Преимущество стандарта V.42bis заключается в его гибкости – он согласует между модемами значения параметров: максимального размера кодового слова, их общего числа, размер символа, числа символов в алфавите и максимальную длину последовательности. Кроме того, алгоритм осуществляет мониторинг входного и выходного потока данных для определения эффективности сжатия. Если сжатия не происходит или оно невозможно (в силу природы передаваемых данных), алгоритм прекращает свою работу. Это свойство обеспечивает лучшие рабочие характеристики при передаче файлов, которые уже были сжаты или которые не поддаются сжатию.

Однако модем работает постоянно с непрерывным динамическим потоком данных, и, к сожалению, сжатию подвергаются лишь отдельные, относительно небольшие и независимые фрагменты потока, что не позволяет достичь столь же высокой степени сжатия, как в архиваторах. Например, текст на русском языке большинством архиваторов сжимается в 4–5 раз, в то время как реальная эффективность лучших модемных протоколов сжатия не превышает 2–3, а более высокая степень достигается лишь при передаче повторяющихся серий (таблиц, неупакованных баз данных с высокой избыточностью и т.п.).

Для наиболее эффективного сжатия данных необходимо учитывать информационное содержание потока, относящегося к разнородным видам информации.

Таким образом, одним из путей развития алгоритмов сжатия информационных потоков можно считать разработку дополнительных адаптивных алгоритмов сжатия представительного уровня сети, учитывающих статистические свойства информации.

Адаптивный алгоритм сжатия данных. Принцип работы предлагаемого алгоритма (рис. 3) заключается в следующем:

- на первом этапе происходит чтение заголовка файла передаваемых данных с целью определения его формата;

- на втором этапе происходит анализ формата файла и выбор соответствующего алгоритма сжатия: для текстовых файлов – алгоритм "LZW", для звуковых – алгоритм "mp3", для графических файлов – алгоритм сжатия в зависимости от статических свойств изображения;

- на третьем этапе работы алгоритма происходит сжатие информации по выбранному алгоритму, для неизвестных форматов файлов – пе-

реход к следующему этапу;

– на четвертом этапе происходит формирование выходного потока данных с присвоением каждому сжатому файлу данных маркера с указанием его типа и параметров;

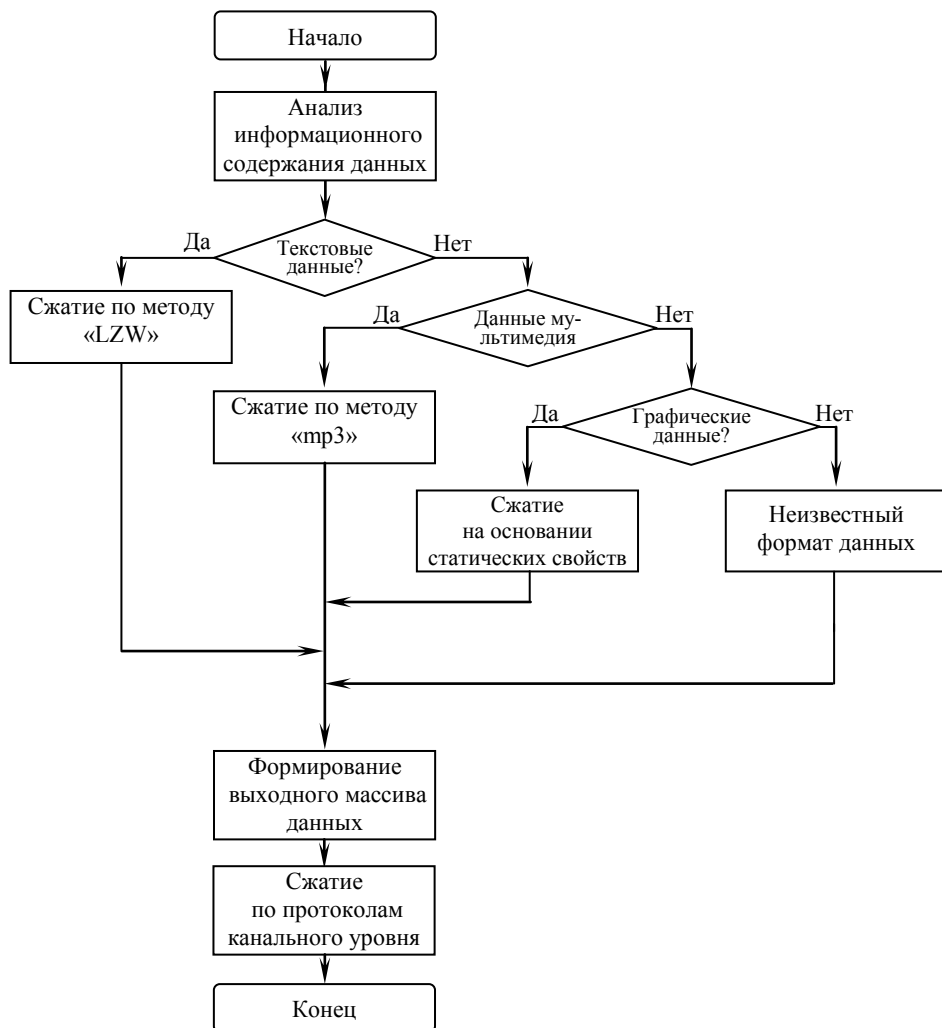


Рис. 3. Адаптивный алгоритм сжатия данных

– на последнем этапе работы сформировавшийся поток данных поддается обработке в зависимости от типа передающей аппаратуры канальным протоколом сжатия, описанным выше.

Алгоритм восстановления сжатой информации после передачи заключается в обратном выполнении указанных действий:

- анализ маркера сжатого файла информации;
- выбор соответствующего алгоритма восстановления информации;
- восстановление сжатого файла, формирование начального потока информации.

Выводы. Таким образом, применение данного адаптивного подхода к сжатию потоков информации в компьютерных сетях, несмотря на увеличение времени сжатия и восстановления информации, позволит уменьшить объем передаваемой информации от 2-х до 10 раз в зависимости от особенностей информационного потока.

Одним из путей дальнейшего развития адаптивного подхода является анализ нестационарных свойств источника и обеспечение согласования выходного потока данных с характеристиками сети передачи данных.

ЛИТЕРАТУРА

1. *Ватолин Д.С. Алгоритмы сжатия изображений. Методическое пособие.* – М.: Издательство МГУ, 1999. – 9 с.
2. *Королёв А.В., Малахов С.В., Линник Н.Ф. Модифицированное зональное сжатие изображений при частичном устранении фазовых составляющих спектра // Системы обработки информации.* – Х.: ХВУ. – 2001. – Вып. 5(15). – С. 176 – 180.
3. *Кульгин М.В. Технологии корпоративных сетей. Энциклопедия.* – С.-Пб.: Питер, 2000. – 704 с.
4. *Климов А.С. Форматы графических файлов.* – С.-Пб.: ДиаСофт, 1995. – 365 с.
5. *Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы.* – С.-Пб.: Питер, 1999. – 150 с.
6. *Сван Т. Форматы файлов Windows.* – М.: Бином, 1995. – 174 с.

Поступила 26.05.2003

РУБАН Игорь Викторович, канд. техн. наук, доцент, зам. нач. кафедры ХВУ. В 1990 году окончил ХВВКИУ РВ. Область научных интересов – обработка информации в вычислительных сетях и АСУ.

КОЛМЫКОВ Максим Николаевич, адъюнкт ХВУ. В 2000 году окончил ХВУ. Область научных интересов – обработка информации в вычислительных сетях и АСУ.

РЕЗУНЕНКО Андрей Александрович, преподаватель Полтавского военного института связи. В 2001 году окончил ХВУ. Область научных интересов – обработка информации в вычислительных сетях и АСУ.