ТЕОРИЯ КОДИРОВАНИЯ КАК ОПТИМИЗАЦИОННАЯ ПРОБЛЕМА ДЕКОДИРОВАНИЯ ВБЛИЗИ ГРАНИЦЫ ШЕННОНА

член-корр. РАН Зубарев Ю.Б.[1], проф. Золотарёв В.В.[2], проф. Овечкин Г.В.[3]
[1]Московский научно-исследовательский телевизионный институт
[2]Институт космических исследований РАН
[3]Рязанский государственный радиотехнический университет

В связи с успешным решением Оптимизационной Теорией (ОТ) помехоустойчивого кодирования проблемы Шеннона анализируется ситуация, сложившаяся в прикладных исследованиях главной отрасли информатики – теории кодирования. Обсуждаются нерешенные проблемы «классической» алгебраической теории и рассмотрены результаты для основных кластеров алгоритмов декодирования, построенных на основе ОТ. Отмечены очередные задачи развития ОТ и предлагаются методы их решения. Обсуждается ценность основных парадигм ОТ. Подчёркиваются достоинства символьных кодов и дивергентного кодирования. Отмечается особая группа алгоритмов декодирования с прямым контролем метрики (ДПКМ). Названы основные пути развития теории кодирования на ближайшие годы и на перспективу.

Введение. Быстрое и успешное решение в России прикладных и теоретических проблем в области создания простых высокоэффективных алгоритмов помехоустойчивого кодирования цифровых данных в каналах с шумами, о чём наше научное сообщество регулярно информируется на конференциях и в научной периодике, а также регулярное издание монографий и справочников по этой теме, свидетельствует о высоком потенциале российской информатики в этой отрасли знаний и технологий. Более 600 блоков справочно-методических данных на сетевых ресурсах ИКИ РАН и РГРТУ [1] оперативно отражают динамику обширных исследований научной школы Оптимизационной Теории (ОТ) кодирования, число публикаций которой давно превысило 400 работ, часть из которых можно найти, например, в [2-7]. Краткому изложению проблем в области теории кодирования и посвящена данная работа.

Состояние классической теории кодирования. Через 70 лет после публикации великой статьи К. Шеннона [8], поставившей перед наукой задачу высокодостоверной цифровой связи по каналам с шумами, успехи теории кодирования, которой пришлось решать эту проблему, можно считать и значимыми, и, одновременно, весьма ограниченными. Её важнейшие результаты 60-летней давности в алгебраических кодах стали этапными достижениями для двоичных симметричных каналов (ДСК) без памяти, хотя они оказались очень далеки от тех возможностей кодирования, которые были предсказаны Шенноном. Поэтому неудивительно, что при сложности N декодирования этих кодов, меньшей даже, чем $N{\sim}n^2$, где n – длина используемого кода (а сложность сразу стала одним из главных оценочных критериев применимости методов, исправляющих ошибки!), затем, однако, наступил период повсеместного применения алгоритма Витерби (АВ), сложность которого росла с длиной кодера K экспоненциально, но его помехоустойчивость даже с короткими кодами в гауссовских каналах была несравненно более высокой, чем у алгебраических кодов.

После появления недвоичных кодов Рида-Соломона (РС), которые быстро стали очень полезными в очень многих технических приложениях, оказалось, что за 60 последующих лет никто так и не придумал каких-либо других недвоичных кодов с простым декодированием и хорошей корректирующей способностью. А в реальности до сих пор можно использовать только весьма короткие коды РС, которые из-за этого неэффективны и работают только при весьма малом входном уровне шума. Но их декодеры тоже имеют значительную сложность ~n^2. Наверное, лучшим примером практического применения алгебраических кодов сейчас является только каскадная схема АВ с кодом РС. Но достоинства этой схемы определяются, в основном, используемым свёрточным кодом и декодером АВ.

Дальнейшая судьба алгебраической теории незавидна. Двоичные коды этого класса не смогли работать в гауссовских каналах так же легко, как АВ, и не исправляли ошибок за пределами половины минимального кодового расстояния d. Более того, не нашлось даже таких способов их обработки (декодирования), сложность которых росла бы с увеличением длины кодов лишь линейно, т.е. имела бы порядок $N{\sim}n$.

Подчеркнём, что все алгебраические коды работали только в области вероятностей ошибок канала, которые были весьма малы относительно тех их значений, которые допускала теория. Это и стало концом прикладной алгебраической теории кодирования.

Преодоление первого кризиса. Высокие характеристики достоверности алгоритма АВ в двоичных каналах быстро вывели технику декодирования на уровень, который был вполне достаточен для разных цифровых систем в течение нескольких десятилетий. Но он крайне сложен из-за того, что должен помнить все возможные решения, число которых растёт с длиной кода экспоненциально. Напомним, что АВ является оптимальным декодером (ОД), минимизирующим вероятность ошибки своих решений, сравнивая расстояния между принятым сообщением до всех потенциальных решений, выбирая затем самое близкое из них.

Первая успешная попытка преодоления длительного масштабного кризиса в теории кодирования состояла в том, чтобы ориентироваться на эффективность АВ, но не хранить все решения в декодере, а выбрать сначала определённым образом некоторую исходную гипотезу-решение для нового алгоритма. А затем было предложено последовательно находить только более близкие к принятому вектору решения, которые, возможно, иногда будут приводить к решениям оптимального декодера (ОД). Интересно, что даже после публикации полвека назад этой привлекательной идеи, уменьшающей сложность таких декодеров сразу до линейной, пропорциональной $N \sim n$, т.е. до минимально возможной, реализовала её только научная школа ОТ. Эти новые алгоритмы для двоичных каналов были названы многопороговыми декодерами (МПД). Они ещё 40 лет назад сразу показали при лишь линейной сложности высочайшую достоверность и фактическую оптимальность своих решений при уровнях шума, совершенно недоступных для алгебраических методов. А затем были созданы и методы поиска таких кодов, которые быстро довели достоверность решений МПД при их использовании до уровня ОД даже при кодовых скоростях, очень близких к границе Шеннона.

Решение проблем недвоичных и других кодов. Для осознания проблемы простого декодирования недвоичных кодов после открытия и патентования двоичных МПД потребовалось ещё 8 лет, прежде чем специальные коды, названные символьными, участники школы ОТ научились оптимально декодировать с минимально возможной, т.е. линейной от длины кодов сложностью. Результаты были столь необычными, что, как и для двоичных кодов, для ОТ оказалось полезным оформить целый ряд патентов и на символьные декодеры [7]. Длины кодов n и размер алфавита q для символьных кодов, в отличие от кодов РС, абсолютно независимы, что тоже оказалось крайне удобным для многих технических приложений. Совпадение решений символьных декодеров с результатами для ОД даже при большом уровне шума, а также линейная сложность символьных МПД навсегда утверждают лидерство этих новых алгоритмов среди недвоичных алгоритмов. У них уже более 30 лет нет вообще никаких конкурентов.

Ещё один крайне проблемный аспект теории кодирования таков, что для недвоичных кодов в реальности вообще невозможно создать АВ. Даже для простейшего однобайтового очень короткого и поэтому малоэффективного кода с длиной кодера $K=5$, у которого $q=256$, сложность недвоичного АВ будет порядка $N \sim q^K = 10^{12}$, что, конечно, реализовать нельзя. Тем самым исключительная важность полного решения проблемы оптимального декодирования недвоичных кодов на основе символьных МПД, да ещё с линейной сложностью (!), становится особенно очевидной и крайне актуальной.

Отметим также, хотя результаты ОТ давно опубликованы, в том числе и на английском, до сих пор нет никаких данных даже о повторении наших результатов где-либо вообще. Уникальность и пока ещё полная недоступность достижений ОТ другим группам исследователей иллюстрируется рис. 1. Это объясняется тем, что для исследований МПД и других алгоритмов необходимо использование высокоинтеллектуального программного обеспечения системного стиля и высочайшего уровня.

Отметим также, что нами запатентован ещё и простейший декодер для стирающих каналов с минимальной сложностью, восстанавливающий стёртые символы непосредственно вблизи пропускной способности канала [1, 7].

Очень важно, что методы ОТ, в отличие от турбо и низкоплотностных кодов, не используют вычислений с действительными числами. Это ещё более выделяет технологичность и эффективность алгоритмов МПД, что особенно важно при аппаратной реализации декодеров. Но работа только с фиксированной точкой, т.е. только с небольшими целыми числами, заметно ускоряет и работу алгоритмов, реализованных программно.

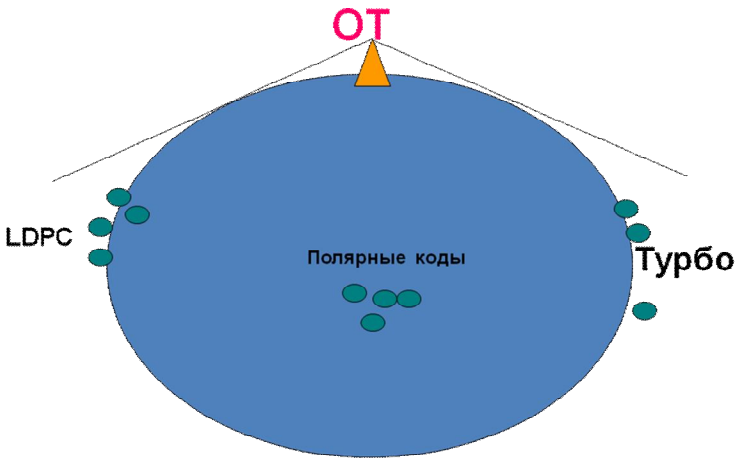

Рисунок 1

Наконец, очень значимое достижение ОТ состоит в том, что мы создали и запатентовали особую версию АВ для блоковых кодов (БАВ) с такой же сложностью N, как и у самого знаменитого в мире классического свёрточного АВ, т.е. с $N{\sim}2^K$, где K – длина кодирующего регистра, тогда как до самого последнего времени наши «теоретики» гордились блоковыми АВ с $N{\sim}2^{2K}$ [9]. Можно для сравнения указать, что, как известно [7], для одного из проектов NASA использовался АВ с K=15. Наш блоковый АВ был бы для этого кода в ~16'000 раз проще той странной «игрушки» теоретиков, которая по-прежнему предлагается в [9] для обучения студентов.

Главные результаты ОТ. Краеугольным камнем всей ОТ является Основная Теорема многопорогового декодирования (ОТМПД), простейшая по форме, но наполненная глубочайшим системно-философским смыслом. Согласно этой Теореме при линейной сложности этого алгоритма декодирования решения МПД стремятся к решениям ОД. При этом в ОТ сразу указывается, что МПД не ОД и после некоторого процесса улучшения своих решений, алгоритм может остановиться раньше, чем достигнет оптимального решения. Школа ОТ в течение десятилетий искала и создала много методов, которые позволили строить коды, при использовании которых МПД разных типов даже при большом уровне шума для всех основных классических моделей каналов связи достигают именно решений ОД.

Все основные результаты ОТ базируются на решении оптимизационных задач различных типов. Полезно отметить и то, что граница Шеннона принципиально недостижима, поскольку является, как и скорость света для материальных тел, абсолютно упругой. А это приводит к тому, что между пропускной способностью каналов и рабочей областью МПД всегда есть некоторый промежуток, который, по опыту прошедших лет, усилиями участников школы ОТ постоянно уменьшается. Ситуация с улучшением характеристик МПД представлена на рис. 2, иллюстрирующем стремление рабочих характеристик МПД к границе Шеннона, определяющей в данном случае предельный уровень шума, при котором теоретически еще возможна сколь угодно надежная передача данных при заданных параметрах кодера и канала. Изложенные свойства МПД подтверждаются разнообразными результатами в [1, 6, 7].

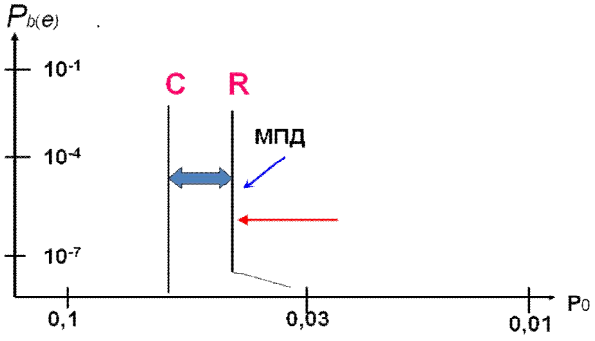

Рисунок 2. О стремлении к границе Шеннона

Основные технологии ОТ. Развитие прикладных методов кодирования, которые позволили полностью решить великую проблему Шеннона и перевести задачу из чисто научной в технологии на базе теорий глобального поиска, определяется главной теоремой ОТМПД, доказанной в ОТ для всех классических каналов, рассматривавшихся ранее в теории кодирования, а также уже достаточно совершенной теорией размножения ошибок (РО) в мажоритарных декодерах, которая создала мощные технологии построения таких мажоритарных кодов, которые позволили алгоритмам МПД успешно работать вблизи границы Шеннона. Но для получения действительно высокой достоверности решений при теоретически минимально возможной сложности МПД и непосредственной близости кодовой скорости R к пропускной способности цифрового канала C, т.е. при $R \lesssim C$, в ОТ были созданы и успешно использованы десятки способов, которые были запатентованы, а также реализованы в аппаратном виде или программно. Они создали совокупность множества простых методик, обеспечивших реальное решение проблемы Шеннона.

В первую очередь, это дивергентное кодирование, позволяющее некаскадными методами последовательно увеличивать кодовое расстояние применяемых кодов. Далее это выделение группы декодеров с прямым контролем метрики (ДПКМ), также сильно продвинувших характеристики всех лучших алгоритмов в сторону ещё большого уровня шума. Велика здесь и роль параллельного каскадирования, первооткрывателем которого является наша школа [10]. Важны и наши новые подходы к традиционным схемам последовательного каскадирования, поскольку характеристики МПД с оптимальным декодированием принципиально меняют вклад в общую эффективность от обоих кодов, входящих в систему каскадирования [1, 5-7].

О соотношении реального и идеального в науке. Проблема соотношения возможностей теории и эксперимента особенно обострилась в 80-х годах в связи с бурным развитием вычислительной техники. Часть публикаций на эту тему была представлена даже на портале РАН [11]. Приведём из этой статьи лишь фразу: «....противостояние компьютерного моделирования и теории, основанной на математических методах, – Болезнь Века». Она и стала точным диагнозом глубокого кризиса «классической» теории кодирования.

Теоретики 60-х годов вполне обходились простыми формулами, например, биномиального распределения, характеризовавшими достоверность весьма малоэффективных алгебраических алгоритмов декодирования. Но характеристики АВ и других методов, в том числе МПД, уже нельзя было посчитать точно. Ими занялись другие специалисты, активно использующие компьютерное моделирование. Но теоретики остались «внутри» алгебраической теории, где можно было многое просчитать «в уме» или аналитически. И даже когда появились программы моделирования для студентов, которые позволяли проводить самые простые эксперименты в области теории кодирования, ощущения необходимости широкого использования именно полномасштабного моделирования на компьютерах всех алгоритмов декодирования у «теоретиков» не возникали. Они вполне обходились «ансамблями кодов», для которых научились выполнять весьма заумные оценки, а потом защищали их в таких же «теоретических» советах. И тогда у них всё остановилось. Новые достоверные проверяемые алгоритмы у «теоретиков» больше вообще не появлялись.

Но именно осознание коллективом школы ОТ ещё в ~1975 г. единства теории и эксперимента в теории кодирования позволило этой школе создать ряд специализированных программных систем для исследований прикладных проблем теории кодирования. Они применялись для моделирования различных цифровых систем с развитыми методами обработки и настройки параметров кодов и декодеров. Такой подход создал мощный синергетический эффект ускорения исследований. Теория без эксперимента крайне слаба, а оторванный от теории эксперимент почти бесполезен и, нередко, просто ошибочен. Вполне уместно указать и на то, что у «теоретиков» кодирования, «достижения» которых очень часто проверить невозможно, доля ошибок весьма велика [11]. В школе ОТ используются много способов управления изощрёнными экспериментами на основе порой даже несколько экзотической, но одновременно и очень логичной теории. А такая взаимопроверка и поддержка оказывается очень продуктивной и крайне полезной и для моделирования, и для теории.

О текущем уровне «прочей» теории кодирования. Что же касается довольно больших групп наших «теоретиков», не использующих моделирование, приходится согласиться с тем, что, видимо, никто из них не владеет современными методами программирования,

позволяющими создавать полноценные модели декодеров и точно оценивать их достоверность, помехоустойчивость и сложность! Хлеб программистов очень тяжкий, они не относятся к привилегированным профессиям, но их, видимо, полное отсутствие в командах, «играющих» с кодами, теперь уже навсегда обнулило результативность «теоретиков». В реальности же это привело к тому, что все такие группы, которые заняты «работой» в «чистой» теории кодирования, не предложили за многие последние десятилетия вообще ни одного конкретного декодера. А алгоритмы для сотен «ансамблей кодов», успешно защищённых и в качестве докторских диссертаций ещё с 90-х годов того века, являются скучными фикциями, т.к. декодеры можно создавать только для конкретных кодов. И сейчас многие докторские диссертации объявляют о своих «выдающихся» результатах для совершенно неэффективных и весьма непростых декодеров со сложностью $N{\sim}n^2$, тогда как ещё в 1981 г. были опубликованы данные об МПД алгоритмах со сложностью $N{\sim}n$ и с характеристиками оптимальных декодеров [12, 13], а 15 лет назад все эти сведения уже были помещены и в соответствующий справочник по кодированию [6]. К этому же удивительному списку «результатов» можно отнести целый ряд докторских диссертаций, главным «достижением» которых было снижение в 3-10 раз экспоненты сложности декодирования каких-то кодов по сравнению с полным перебором. Но это практически всегда соответствует даже для не очень длинных кодов, например, с $n{\lesssim}10^4$, сложности декодера N, превышающей число атомов во Вселенной [9]. Понятно, что это очень «практично и ценно»!

Отдельно следует также подчеркнуть (см. рис. 1), что и полярные коды тоже не стали хоть каким-то выходом из кризиса «той» прежней теории кодирования. Мы рассмотрели их возможности в [14]. За десятилетие развития ещё никто и нигде не предъявил хотя бы один полный комплект реальных характеристик таких декодеров. Здесь мы пока отметим, что в докторской диссертации [15] по этой тематике большинство упомянутых там алгоритмов имеют сложность $N{\sim}n^2$. А в кандидатской работе [16] для этих кодов обсуждаются декодеры со сложностью более $N{\sim}n^3$.

Наконец, отметим, что студентов во всех ВУЗах РФ, включая самые элитные, через 30 лет после публикации полной теории ОТ и, повторим, издания общедоступного справочника [6] учат, в основном, по алгебраическим книжкам и программам курсов, в которых иногда нет даже алгоритма Витерби. Это тоже абсолютно необходимо срочно исправить.

Общая ситуация в теории кодирования. К настоящему времени в нашей прикладной ОТ рассмотрены все основные методы кодирования, пригодные для самого широкого использования. Они работают в непосредственной близости от границы Шеннона при линейной сложности алгоритмов МПД, достигающих решений оптимальных декодеров. Кроме того, МПД декодеры имеют теоретически максимально возможные скорости аппаратного декодирования, что демонстрирует разработанный в ИКИ РАН МПД декодер для свёрточного кода, показанный на рис. 3. На базе ПЛИС ALTERA создан декодер на скорость более 1 Гбит/с [2, 3, 7]. Программные версии МПД работают на скоростях от десятков Кбит/с до Мбит/с [1, 7].

Рисунок 3. МПД для космоса, оптических каналов и флеш-памяти

Кроме того, на основе ОТ построены быстрые декодеры для оптических линий, равных которым нет [1]. Разработаны также несколько очень простых алгоритмов декодирования с характеристиками ОД для флеш-памяти, для чего потребовалось на много порядков повысить достоверность результатов декодирования [1, 3, 7].

Заключение. В монографиях [2, 3, 7] и на ресурсах [1] в разнообразных форматах широко представлена Оптимизационная Теория, которая полностью решила проблему Шеннона – задачу простого высокодостоверного декодирования вблизи пропускной способности цифровых каналов. Сопоставимых с методами ОТ по полной совокупности параметров достоверность-помехоустойчивость-сложность других алгоритмов декодирования с подтверждёнными признанными характеристиками пока нет.

Дальнейшее развитие ОТ будет связано с различными системами многопозиционных сигналов [2, 4, 7], а также с более сложно организованными системами и сетями связи. Научная школа ОТ всегда готова всесторонне поддержать перспективные разработки других организаций и инициативных специалистов, которые будут развивать простые действительно высокоэффективные алгоритмы декодирования различных классов и решать другие важные смежные прикладные задачи.

Издание монографий [2, 6, 7] осуществлялось при поддержке РФФИ.

Литература

1. Ресурсы www.mtdbest.iki.rssi.ru и www.mtdbest.ru .

2. Золотарёв В.В., Зубарев Ю.Б., Овечкин Г.В. Многопороговые декодеры и оптимизационная теория кодирования // Под редакцией академика РАН В.К. Левина. М., Горячая линия – Телеком, 2012, 238 с.

3. Zolotarev V., Zubarev Y., Ovechkin G. Optimization Coding Theory and Multithreshold Algorithms // Geneva, ITU, 2015, 159 p.

(e-book reference: http://www.itu.int/pub/S-GEN-OCTMA-2015).

4. Зубарев Ю.Б., Овечкин Г.В. Помехоустойчивое кодирование в цифровых системах передачи данных // Электросвязь, 2008, №12, с.58-61.

5. Золотарёв В.В., Овечкин Г.В. Применение многопороговых методов декодирования помехоустойчивых кодов в высокоскоростных системах передачи данных // Электросвязь. М., 2014, №12, с.10-14.

6. Золотарёв В.В., Овечкин Г.В. Помехоустойчивое кодирование. Методы и алгоритмы. Справочник. Под научной редакцией члена-корреспондента РАН Ю.Б. Зубарева. М., Горячая линия – Телеком, 2004, 126 с.

7. Золотарёв В.В. Теория кодирования как задача поиска глобального экстремума. Под научной редакцией академика РАН Н.А. Кузнецова // М., Горячая линия - Телеком, 2018, 220 с.

8. Shannon C. A Mathematical Theory of Communication // Bell System Technical Journal. Vol. 27 (July and October 1948). P.379–423 and 623–656. (Шеннон К.Э. Математическая теория связи // В сб.: Работы по теории информации и кибернетике. М.: Иностранная литература, 1963)

9. Кудряшов Б.Д. Основы теории кодирования. СПб.: БХВ-Санкт-Петербург, 2016, 393с.

10. Золотарёв В.В. Параллельное кодирование в каналах СПД. В сб.: «Вопросы кибернетики», ВК-120, АН СССР, Научный совет по комплексной проблеме «Кибернетика», М., 1986, с.33-37.

11. Магаршак Ю. Число, возведенное в абсолют // Независимая газета, 09.09.2009 г.

12. Самойленко С.И., Давыдов А.А., Золотарёв В.В., Третьякова Е.Л. Вычислительные сети. // М., Наука, 1981, 278 с.

13. Золотарёв В.В. Эффективные многопороговые алгоритмы декодирования // Научный совет по комплексной проблеме «Кибернетика» АН СССР. Препринт, М.,1981, 76 с.

14. Золотарёв В.В., Овечкин Г.В. О сопоставлении новых методов помехоустойчивого кодирования // Доклады 18-й Международной конференции «Цифровая обработка сигналов и её применение», М., 2016, Т.1, с.59-64.

15. Трифонов П.В. Методы построения и декодирования многочленных кодов // Докторская диссертация, СПб, 2018.

16. Милославская В.Д. Методы построения и декодирования полярных кодов // Кандидатская диссертация, СПб, 2014.

Abstract

Due to the successful solution by the Optimization Theory (OT) of error-correcting coding Shannon's problem, the situation is analyzed in applied researches in the main branch of Informatics - coding theory. The unsolved problems of «classic» algebraic theory are discussed and the results for the main clusters of decoding algorithms constructed on the OT basis are considered. The next tasks of OT development are noted and methods of their solutions are offered. The value of the main OT paradigms is discussed. Emphasizes the dignity of symbolic codes and divergent coding are made. A special group of decoding algorithms with direct metric control (DDMC) is noted. The main ways of development of the theory of coding in the coming years and in the future are named.