

УДК 621.391

Математическая теория резервирования на основе модели формального нейрона

Олег Намичейшвили, Арчил Элизбарашвили

Резюме

Статья вводит в круг основных идей и методов математической теории порогового резервирования цифровой вычислительной техники на основе модели формального (искусственного) нейрона и ставит целью способствовать дальнейшему развитию этого направления теории надёжности, переживающей пору расцвета. Содержание статьи концентрируется на следующих основных результатах:

- математическая модель и соответствующие алгоритмы для оценки вероятности ошибки порогового восстанавливающего органа, а также минимальной верхней границы этой величины;

- асимптотическая оценка вероятности ошибки мажоритарного элемента;

- энтропийный подход к назначению оптимальных входных весов порогового решающего органа и доказательство его адекватности с байесовским подходом;

- получение в явной форме выражений для весов, доставляющих максимум обобщённому махаланобисову расстоянию и установление существования монотонного преобразования по гиперболическому синусу между ними и весами, вычисляемыми на основе энтропийного подхода;

- установление законов распределения и других вероятностных характеристик входных весов порогового восстанавливающего органа при статистическом разбросе вероятностей ошибок на этих входах;

- доказательство возможности реализации пороговым восстанавливающим органом функции решения с минимальным риском и получение необходимого для этого порога.

Ключевые слова: *надёжность, резервирование, формальный (искусственный) нейрон, кворум, пороговый элемент, решающий орган, вероятность ошибки восстановления сигнала, двоичный (бинарный) сигнал, обобщённое (махаланобисово) расстояние, энтропия случайной величины, байесовский подход, оптимизация, входные веса.*

Введение

Обеспечение надёжности бинарных (двоичных) информационных каналов методами введения избыточности представляется важнейшим аспектом проблемы надёжности. Повышение интереса к этой проблеме в последние годы, вызванное стремлением предотвратить отказы сложных комплексов, используемых в авиации, космической технике, энергетике и транспорте, стимулировало развитие и совершенствование методов введения избыточности для обеспечения работоспособности управляющей сложными объектами вычислительной техники.

Известные к настоящему моменту методы введения избыточности в вычислительные устройства основаны на применении

- органов голосования и устройств, использующих статистическую теорию принятия решений;
- схем релейного типа со специальной архитектурой;
- теории кодирования;
- схем со сложной внутренней структурой соединений;
- логической стабилизации.

Тем не менее проблема повышения надёжности логических структур на основе модели формального (искусственного) нейрона (порогового элемента) всё ещё остаётся малоизученной и актуальной. В частности дополнительного рассмотрения требует вопрос о вероятности ошибки на выходе порогового элемента и её асимптотическом поведении при увеличении числа резервных каналов. Крайне необходимо выявить различные подходы к вопросу оптимального назначения весов для порогового решающего (восстанавливающего) органа и провести их сравнительный анализ. Поскольку оптимальные веса, по-видимому, зависят от значений вероятностей ошибок соответствующих каналов, а эти вероятности оцениваются в результате независимых испытаний частотой появления ошибки, то следует установить законы распределения статистических характеристик весов. При трактовке мажоритарного восстанавливающего органа как частного случая порогового неизбежно встаёт далеко не праздная при микросхемной технологии исполнения задача получения явной зависимости вероятности ошибки на выходе мажоритарного элемента от числа n резервных каналов при $n \rightarrow \infty$. Представляет также интерес установление возможности описания работы восстанавливающего органа в терминах теории распознавания образов и получение формулы для выбора порога, реализующего функцию решения с минимальным риском.

Поскольку целью настоящей статьи является обобщение т.н. кворумного, или порогового подхода (на основе модели искусственного нейрона) к формированию теории резервирования цифровых систем, уместно дать краткий обзор исследований как по этому вопросу, так и по смежным проблемам.

В качестве основополагающей должна быть названа статья [1] Дж. фон Неймана. Здесь предложены две схемы с избыточностью. В одной из них избыточность растёт по показательному закону от т.н. «логической глубины». Такую схему, даже при современном уровне развития микросхемотехники, трудно использовать. В другой же схеме, названной многократной, или мультипликационной, для передачи сигнала применяется система информационных каналов. В системе содержатся несколько независимо вычисленных версий одного и того же двоичного сигнала. После каждой логической операции в схеме используется так называемый восстанавливающий орган, функция которого сводится к обработке избыточной информации на входе системы информационных каналов для получения более надёжной информации на её выходе. Дж. фон Нейманом были предложены логические схемы, выполняющие эту функцию восстанавливающего органа, в том числе и

т.н. мажоритарные элементы. В последних выходной сигнал определяется путём «голосования» входных сигналов и принятия решения по принципу простого большинства.

Основы теории резервирования заложены Ш.Л. Бебиашвили в статье [2], где независимо от Дж. фон Неймана и почти одновременно с ним доказана возможность построения надёжных систем из ненадёжных компонентов.

Важное значение в теории проектирования избыточных структур отводится работе [3] Э.Ф. Мура и К.Э. Шеннона. Полученные в ней результаты касаются релейных схем, особенностью которых является двустороннее распространение сигналов, между тем как в логических вентильных устройствах информация пропускается лишь в одном направлении. Тем не менее результаты работы [3] могут быть использованы и при проектировании избыточных логических схем.

Концепция мажоритарной логики и задача повышения на её основе надёжности цифровых систем обсуждается в работе [4] Р. Миллера, а также в исследовании [5] Р. Вассермана, У. Броуна и Дж. Тьернея. Эти работы развивают лишь те идеи фон Неймана, которые касаются принципа устроения и правила большинства.

Качественно новый этап связан с именем Дж. Трайона, предложившего в работах [6] и [7] введение избыточности по логике учетверения. В методе Дж. Трайона не требуются восстанавливающие органы. Исправление всех одиночных и некоторых многократных ошибок вследствие отказов осуществляется учетверением логических элементов и входных сигналов при жёсткой организации чередования слоёв, составленных из элементов определённого типа. Идея метода основывается на том простом факте, что ошибки различного типа неодинаково воздействуют на элементы различного типа (принцип ассиметрии). Поэтому ошибка, возникшая в слое элементов одного типа, может быть исправлена в следующем слое, составленном из элементов другого, строго определённого типа. Созданный первоначально применительно к реализации в базисе И, ИЛИ, НЕ, метод был распространён Дж. Трайоном также на базис И-НЕ и ИЛИ-НЕ, где эффективность этой идеи особенно велика.

Новый подход, основанный на использовании статистической теории принятия решений, был развит У. Пирсом в исследованиях, выполненных в 1960-1962 годах в лаборатории электроники твёрдого тела Стэнфордского университета. Их содержание отражено в работах [8-11] и подытожено в книге [12].

Развивая идеи Дж. фон Неймана, У. Пирс предложил, в случае, когда надёжность резервируемых информационных каналов неодинакова, использовать т.н. пороговые, или кворумные восстанавливающие элементы, фактически являющиеся моделью формального (искусственного) нейрона. В пороговых, как и в мажоритарных элементах, выходной сигнал восстанавливается путём «голосования» входных сигналов, но при этом веса голосов не обязательно одинаковы. Как известно, в пороговом элементе сигнал на выходе равен логической единице, если сумма взвешенных значений сигналов на входах превышает некоторый порог (кворум), и – логическому нулю в противном случае. Естественно, что более надёжному каналу соответствует больший вес входного сигнала в восстанавливающем органе. Разумеется, не представляет особых трудностей подбор резервных каналов с практически одинаковым исходным уровнем надёжности, однако с течением времени вероятности ошибок двоичных каналов изменяются неодинаково. Поэтому возникает необходимость в соответствующем изменении весов голосов. С этой целью У. Пирс предложил использовать адаптирующиеся восстанавливающие органы, в которых веса устанавливаются в соответствии с изменением вероятности ошибки информационного канала резервной группы. Эту вероятность У. Пирс предлагает определять посредством измерения частоты совпадения сигнала на выходе соответствующего информационного канала либо со значением сигнала на выходе восстанавливающего органа (с так называемым «решением»), либо с правильным откликом, формируемым вне восстанавливающего органа

каким-либо другим способом. Применение адаптирующегося восстанавливающего органа обеспечивает принятие правильного решения даже при одновременном сбое значительного числа двоичных каналов, составляющих резервную группу.

Расширение У. Пирсом понятия восстанавливающего органа, первоначально введённого Дж. фон Нейманом, послужило важным шагом к постижению того внутреннего единства, которое присуще подходам фон Неймана и Трайона, хотя в трайоновской схеме объём оборудования не растёт экспоненциально от логической глубины и, кроме того, в ней не используются решающие элементы. Это единство проявляется в том, что оба подхода позволяют исправлять ошибки отдельных каналов. В указанном плане важная роль принадлежит и работе [13] О. Лоуэншусса, в которой предложено обобщение конструкции восстанавливающего органа в избыточных автоматах на случай K – значной логики. Следует отметить также работы [14] и [15] У. Манна, посвящённые изучению процессов восстановления в вычислительных системах с избыточностью и методам резервирования логических схем. Однако обобщающим шагом к объединению подходов фон Неймана и Трайона следует признать теорию логического проектирования избыточных структур, изложенную в монографии [12] У. Пирса. Им предлагается метод введения избыточности, состоящий в том, что каждый из элементов неизбыточного варианта проектируемой схемы с избыточностью выполняется как некоторый модифицированный элемент, получающий на свои входы в n (величина избыточности) раз большее число сигналов. Применение такого подхода позволяет исправлять в схеме некоторое число ошибок, зависящее от n . Следовательно, исправление ошибок происходит в тех же самых цепях, которые выполняют логические функции. В схеме увеличивается не число элементов, а их «сложность», которую условно можно оценивать числом входных сигналов, подаваемых на элемент.

В теории логического проектирования избыточных структур У. Пирса возникает проблема организации связи между избыточными логическими элементами. В тех случаях, когда не все избыточные сигналы поступают на каждый избыточный логический элемент, структура связей становится очень сложной. В таких случаях избыточные сигналы в требуемой структуре должны быть перемешаны так, чтобы обеспечить максимальную статистическую независимость между различными избыточными сигналами. Вследствие наличия такой особенности У. Пирс называет свою теорию также логикой переплетений (связей). Предложенный им метод построения избыточных логических схем с исправлением ошибок имеет много общего как со способом Трайона, так и со схемой восстанавливающего органа фон Неймана.

Отдельные проблемы обеспечения надёжности управляющих вычислительных устройств методами введения избыточности в их структуру достаточно подробно освещены в книге [16] В.И. Левина по матричной теории надёжности автоматов.

Направление, связанное с повышением надёжности управляющих вычислительных систем методами избыточного кодирования информации, берёт начало в работе [17] Д. Армстронга. Многие аспекты этой проблемы изложены в монографии [18] Я.А. Хетагурова и Ю.П. Руднева, а также в книге [19] Ю.Л. Сагаловича.

Разработанные в технике связи коды с обнаружением и исправлением ошибок [20] успешно используются с целью повышения надёжности информационных каналов систем управления технических систем. Применение находят и специальные коды, предназначенные для быстрого кодирования и декодирования [21]. Однако, как отмечается в работе [25] Б. Диллона и Ч. Сингха, стоимость таких систем в полтора раза превосходит стоимость обычных систем.

Наконец, среди известных к настоящему времени методов введения избыточности в вычислительные устройства (логические сети) следует назвать открытую Л.М. Вербином логическую стабилизацию [23]. Логической стабилизацией называется синтез такой логической структуры (схемы), которая даёт правильный сигнал на выходе даже в том

случае, когда логические цепи этой структуры не могут быть однозначно описаны некоторой булевой функцией [13, с.18]. Хотя возможность построения логически стабильных схем и доказана, теория синтеза таких структур остаётся в эмбриональном состоянии.

Развитие теории пороговой модели применительно к логическим системам шло после появления работы [1] Дж. фон Неймана по многим направлениям, среди которых можно выделить следующие:

- синтез схем из пороговых и мажоритарных элементов;
- изучение свойств мажоритарных элементов с целью определения требований к ним по заданным для всей системы показателям надёжности;
- оптимальное размещение мажоритарных восстанавливающих органов в системе;
- разработка адаптивного порогового элемента;
- исследование надёжности различных мажоритарных моделей и их сравнительный анализ;
- схемный анализ и проектирование мажоритарных элементов.

Первое из перечисленных направлений представлено в фундаментальных работах [24-33] В.И. Варшавского, М. Дертоузоса, И.Н. Боголюбова Б.Л. Овсиевича, Л.Я Розенблюма, А.Д. Закревского, В.Н. Вавилова, Б.М. Егорова, В.С. Ланцева, В.Г. Тоценко, М. Кона, Р. Линдемана и других авторов. Направление, связанное с изучением свойств мажоритарных элементов с целью определения требований к ним по заданным показателям надёжности системы, вырисовывается в работе [34] Г.Н. Грибанова.

Попытка решения задачи оптимального размещения мажоритарных восстанавливающих органов предпринята в работе [12] У. Пирса для частного случая последовательной структуры.

Разработке адаптивного мажоритарного элемента посвящается работа [35] С.М. Доманицкого. Некоторые важные результаты по адаптивным решающим органам получены и У. Пирсом.

Истоки пятого направления, по-видимому, восходят к работе [36]. Проблемы же схемных решений при проектировании мажоритарных элементов с различными типами логики подробно освещены в книге [37] Н.И. Пакулова, В.Ф. Уханова и П.Н. Чернышева.

Что касается настоящей статьи, в ней исследуются некоторые кворумные модели, даётся их сравнительный анализ и оценивается надёжность на основе концепций теории принятия решений и распознавания образов. Поэтому многие работы [38-54] А.А. Харкевича, Г.С. Себестиана, А.А. Горелика, В.И. Васильева, К.С. Фу, Н.Г. Загоруйко, Ю.Л. Барабаша, Л.А. Растригина, Р.Х. Эринштейна, Р.А. Ховарда и других авторов, посвящённые задачам распознавания и принятия коллективных статистических решений при распознавании, идейно связаны с ней в той или иной степени.

В заключение этого обзора следует отметить, что в настоящее время пути повышения надёжности управляющих вычислительных систем видятся и в развитии мультипроцессорных устройств обработки информации и рациональной организации их параллельной работы [55, 56].

Осуществлён качественный переход к созданию совершенно новых моделей организации вычислительного процесса, о чём наглядно свидетельствуют работы [57-64] Б.В. Анисимова, Е.П. Балашова, В.И. Варшавского, Э.В. Евреинова, Д.А. Пospelова, И.В. Прангишвили, В.Б. Смолова и других исследователей.

В аспекте изучаемых нами задач интересны подходы, развиваемые в работах [65-67] Г.В. Кантария, [68] Г.С. Цирамуа, [69] В.И. Левина, [70] В.И. Потапова, [71] С.М. Доманицкого, [72, 73] А.С. Степанова.

В частности работа Г.С. Цирамуа посвящена проблеме разработки и исследования новых принципов структурной организации, заключающихся в использовании функциональной избыточности многофункциональных элементов в совокупности с адаптацией для повышения надёжности и живучести вычислительных систем.

В основу решения поставленной проблемы положено создание нового класса адаптивно-перестраиваемых вычислительных систем на базе функционально-избыточных элементов, которые в любом интервале времени способны выполнять определённую функцию из заданного множества с целью организации адаптивного самовосстановления системы. Сущность последнего заключается в перераспределении функций между функционально-избыточными элементами в случае утраты отдельными элементами некоторых функциональных возможностей.

Для реализации предложенного принципа автором разработаны методики синтеза и оценки надёжности, а также способы и технические средства адаптивной перестройки по различным законам комбинаторики.

Выше подчёркивалось, что в основе методов синтеза надёжных логических систем, использующих идеи Дж. фон Неймана, Э. Мура и К. Шеннона, лежит принцип накопления информации (путём введения избыточности в форме повторяющихся звеньев ненадёжных модулей) и её последующей обработки восстанавливающим органом. Отмечалось, что подобное накопление информации в комбинационных и последовательностных логических схемах можно осуществлять и с помощью помехоустойчивого кодирования или искусственного усложнения рабочих программ и контрольно-проверочных тестов. В частности исправление ошибок отдельных элементов в комбинационных устройствах с применением самокорректирующихся кодов осуществляется на основе идей, развитых в работах А.Д. Закревского [31, 74] и Т.А. Франциса [75]. Оценка степени улучшения надёжности при применении корректирующихся кодов дана Ю.П. Рудневым [76]. Идея же применения самокорректирующихся кодов в логических схемах с памятью впервые высказана М.А. Гавриловым [77]. В альтернативном подходе, развиваемом В.И. Левиным в его диссертационной работе, эффект ненадёжности интерпретируется как результат взаимодействия случайного сигнала со случайной помехой, а накопленная информация рассматривается как случайная выборка, по которой и требуется восстановить правильный ответ. Основное отличие такой трактовки от классической состоит в широком использовании, наряду с апостериорной, и априорной информации. С учётом сказанного автор развивает методы вероятностного анализа надёжности логического автомата, использующие также идею Г.Н. Церцвадзе [78] о представлении конечного автомата в виде вероятностного. Исследование подобного автомата является сложной проблемой и требует использования матричного исчисления. Для упрощения либо применяется приближённый метод, либо осуществляется дискретное функциональное преобразование.

Рассмотрение конечного автомата в качестве вероятностного позволяет В.И. Левину находить такие характеристики, как вероятность правильного состояния, вероятность правильного выходного сигнала после M изменений входных сигналов, вероятность того, что автомат после M изменений входных сигналов будет находиться в каком-либо из возможных состояний. Но интерес иногда могут представлять и статистические характеристики работы автомата после появления в нём неисправностей. Подходы к решению этой группы задач предложены А.Н. Скляревичем [79].

С.М. Доманицким рассмотрен восстанавливающий орган, в котором веса входных сигналов одинаковы, но при изменении внешних условий может соответствующим образом адаптироваться порог его срабатывания, для чего в восстанавливающий орган вводится

апостериорная информация линейно (плавно) или дискретно (ступенчато). Эта идея находит своё воплощение в восстанавливающем органе с «плавающим» порогом срабатывания. В одном из его вариантов реализуется алгоритм непосредственного изменения порога срабатывания, а в другом – косвенное управление. Автором предложен и соответствующий подход к расчёту избыточных схем, содержащих адаптивные органы голосования указанных типов.

В работе [72] Ю.А. Курочкина и А.Ц. Смирнова проводится интересное сравнение предложенной П.П. Пархоменко и Е.С. Согомояном [80] самопроверяемой избыточной структуры, использующей схемы встроенного контроля, с троированной мажоритарной структурой. При этом, применяя методы расчёта надёжности, разработанные Л.Б. Гройсбергом, М.Д. Линденбаумом [81] и Е.Н. Митичкиным [82], авторам удалось показать, что, если число выходов дискретного устройства больше двух, то при независимой реализации выходных функций троированная мажоритарная структура обладает более высокой надёжностью, чем самопроверяемая избыточная структура. Этим устанавливаются границы области целесообразности применения отмеченных двух классов структур.

В работе [73] развивается рассмотренный С.М. Доманицким [83] подход к расчёту надёжности невозстанавливаемых дискретных устройств с учётом их структуры. В частности авторами исследована связь функциональной и информационной надёжности и получена оценка вероятности отсутствия ошибки в выходном сигнале структурно избыточного устройства без памяти.

Из методов, связанных с введением информационно-временной избыточности в передаваемый сигнал, нами затрагивались выше лишь методы избыточного кодирования. Однако нельзя не отметить особо и достоинства временного резервирования, разработанного Г.Н. Черкесовым [84].

Возможность использования для повышения надёжности вычислительных систем методов временной избыточности связана со структурными особенностями их построения.

Структурную избыточность, заложенную в современные вычислительные системы, вполне удаётся достаточно гибко использовать в различные промежутки времени. При этом очевидно, что должен быть некоторый запас времени для выполнения заданных функций. Этот резерв времени позволяет либо повторить рабочую программу многократно, либо изменить ход её выполнения таким образом, чтобы удалось выявить отказ и устранить его влияние.

Различают два способа [85] организации вычислительных систем с временной избыточностью:

- ограничено время простоя после каждого отказа;
- ограничено лишь суммарное время простоя.

В особый класс можно выделить задачу [86] реализации булевой функции многих переменных на пороговом элементе. В общепринятой терминологии эта задача может рассматриваться как частный случай проблемы воспроизведения булевой вектор-функции на многослойном перцептроне, в котором пороговый элемент (формальный, или искусственный нейрон) представляет собой элементарное звено. Этому направлению в теории искусственных нейронных сетей посвящено большое число работ, опубликованных в течение последних лет. В качестве примера можно привести работу Д.Е. Гурьева [87]. Указанная проблема затрагивается и в обзорной статье [88] Р.П. Липмана.

Таким образом, пороговая логика в литературе рассмотрена достаточно широко и об этом свидетельствуют также интересные результаты [89-101] минувшего десятилетия.

Анализ показывает, что дальнейшее развитие и совершенствование методов повышения надёжности ведёт к более широкому использованию принципов организации биологических систем.

1. Точное выражение вероятности ошибки восстановления

1.1 Описание объекта исследования

Допустим, что двоичный сигнал x , кодируемый, скажем, как $+1$ и -1 , подаётся на n однотипных информационных каналов B_1, B_2, \dots, B_n . Из-за возможных ошибок каналов значение переменной x оказывается вычисленным как x_1, x_2, \dots, x_n . В результате получают n версий для значения предъявленной к распознаванию переменной x . Разумеется, каждая из величин $x_i (i = \overline{1, n})$ также является двоичной переменной, принимающей значения $+1$ и -1 . Эта избыточная информация (в форме n версий для значения переменной x) поступает далее на входы т.н. решающего, или восстанавливающего элемента (органа), модель которого представлена на рисунке 1.

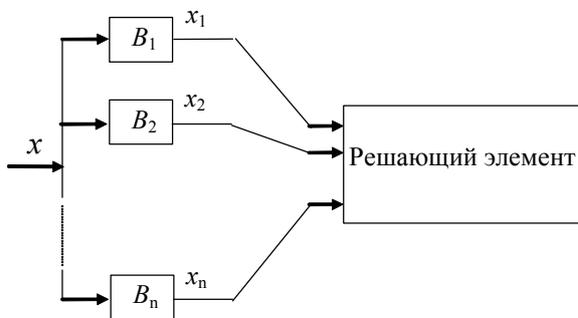


Рис. 1 Модель решающего органа

Как известно, решающим элементом называется устройство, которое по известным двоичным сигналам x_1, x_2, \dots, x_n на n входах определяет двоичный выходной сигнал y , называемый решением. Иначе говоря, решающий элемент представляет собой переключательную схему, реализующую некоторую двоичную функцию y от n двоичных аргументов x_1, x_2, \dots, x_n :

$$y = f(x_1, x_2, \dots, x_n). \quad (1.1.1)$$

Надёжность решающего органа существенно зависит от вида реализуемой решающим элементом функции (1.1.1). Очевидно, что в идеальном случае решение y , принимаемое решающим элементом, должно совпадать с истинным значением двоичной переменной x .

Решающий элемент, реализующий функцию

$$y = \operatorname{sgn} \left(\sum_{i=1}^n x_i \right), \quad (1.1.2)$$

где

$$\operatorname{sgn} z = \begin{cases} -1, & \text{если } z < 0 \\ 0, & \text{если } z = 0, \\ +1, & \text{если } z > 0 \end{cases} \quad (1.1.3)$$

называется мажоритарным. Здесь нулём обозначено неопределённое (безразличное) значение y . Неопределённость выходной переменной y для конкретной входной комбинации x_1, x_2, \dots, x_n означает, что, либо эта комбинация никогда не реализуется (n является нечётным числом в сумме $z = x_1 + x_2 + \dots + x_n$), либо при $z = 0$ решение y не выносится. График функции $y = \operatorname{sgn} z$ представлен на рисунке 2. Стрелки здесь указывают

на то, что острière не принадлежит графику, а точка в начале координат отвечает неопределённости выходной переменной y .

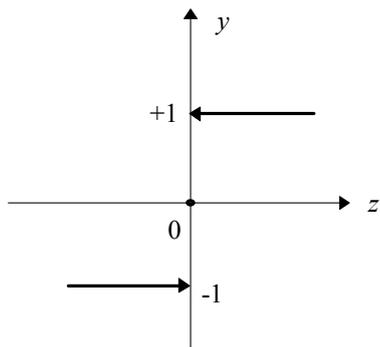


Рис. 2 График функции $y = \text{sgn} z$

Основными компонентами мажоритарного элемента служат суммирующее устройство, с выхода которого снимается сигнал

$$z = \sum_{i=1}^n x_i,$$

и показанный на рисунке 3 нелинейный двухполюсник с характеристикой $y = \text{sgn} z$. Совершенно очевидно, что такой элемент выносит решение y в результате голосования по принципу простого большинства значений сигналов на выходах. Поэтому часто его называют

голосующим. Впервые мажоритарный закон был описан Дж. фон Нейманом в работе [1] и распространён В.И. Варшавским на аналоговые системы с резервированием [25]. Исследованию он подвергался и во многих других аспектах, которых здесь мы не касаемся.

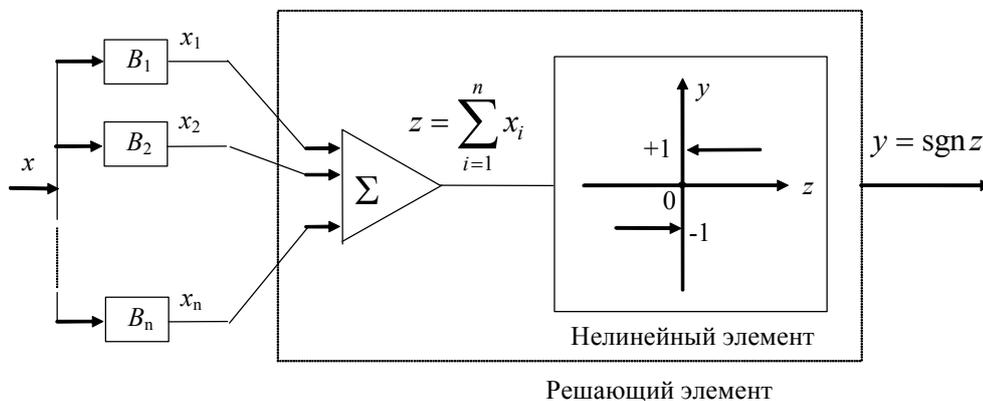


Рис. 3 Модель мажоритарного элемента

Функционирование восстанавливающего органа с мажоритарным решающим элементом не может быть признано удовлетворительным, если вероятности q_1, q_2, \dots, q_n ошибок двоичных каналов B_1, B_2, \dots, B_n различны и, следовательно, каждой из информации x_i , поступающей с выхода двоичного канала B_i на i -й вход решающего элемента, приходится приписывать свой вес $a_i (i = \overline{1, n})$, где a_i - произвольное вещественное число $(-\infty < a_i < +\infty)$. В этом случае решение y на его выходе должно выноситься как результат взвешенного голосования, согласно следующему соотношению:

$$y = \text{sgn} \left(\sum_{i=1}^n a_i x_i - \Theta \right), \tag{1.1.4}$$

где Θ - так называемый порог, или кворум элемента. В силу последнего обстоятельства его часто именуют пороговым, или кворумным, хотя такой элемент с равным правом можно было бы назвать взвешенно голосующим. Нетрудно видеть, что мажоритарный решающий элемент есть пороговый элемент с весовыми коэффициентами $a_i = 1 (i = \overline{1, n})$ и порогом $\Theta = 0$.

Формально допустим, что $\Theta = a_{n+1}$, а $x_{n+1} \equiv -1$. Последнее означает, что имеется некоторый информационный канал B_{n+1} , всегда выдающий сигнал $x_{n+1} \equiv -1$, какой бы сигнал x на его вход ни поступал. Тогда соотношению (1.1.4) можно придать и следующий вид:

$$y = \operatorname{sgn} \left(\sum_{i=1}^{n+1} a_i x_i \right), \quad (1.1.5)$$

Данному соотношению отвечает модель решающего органа, представленная на рис. 4.

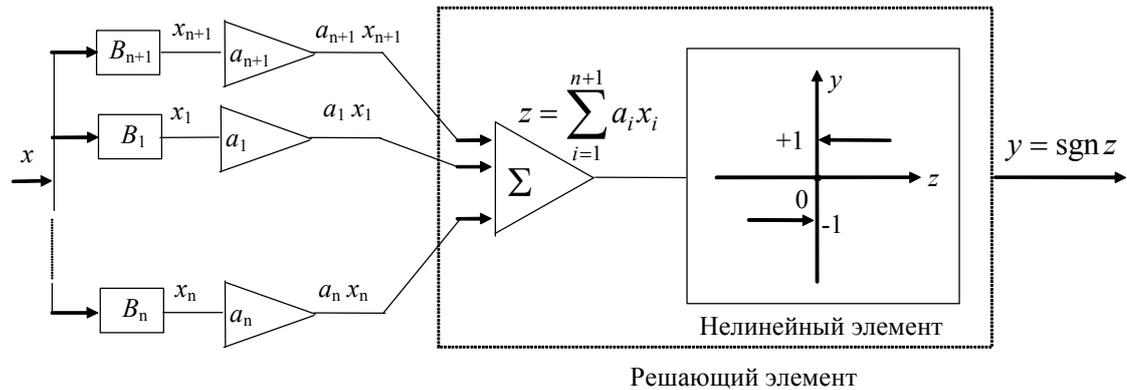


Рис.4 Модель порогового органа

Необходимость в такой модели возникает в связи с тем, что, если в начальный момент времени вероятности ошибок информационных каналов и удаётся подобрать практически одинаковыми, с течением времени в них всё же появляются расхождения.

Пороговая логика освещена довольно широко в литературе [29]. Основные результаты касаются задачи синтеза порогового элемента (иначе, искусственного нейрона), т.е. реализуемости данной переключательной функции n двоичных переменных на одном пороговом элементе (искусственном нейроне) и нахождению необходимых весов и порога, обеспечивающих выполнение этой функции. Решены также задачи синтеза сетей из пороговых элементов, т.е. задачи синтеза искусственных нейронных сетей. Однако в целом упомянутые исследования не затрагивают проблем пороговой логики в аспекте теории принятия оптимальных в некотором смысле решений по восстановлению истинного сигнала при резервировании информационных каналов. Пороговая же модель функционирования решающего органа в системе восстановления правильного сигнала представляет интерес именно с этой точки зрения и служит объектом исследования в данной работе.

1.2 Вероятность ошибки восстановления

Основную цель исследования в настоящем разделе будет составлять получение выражения для вероятности Q того, что принятое пороговым элементом, представленным на рисунке 4, решение y ошибочно, т.е. не совпадает с истинным значением двоичной переменной x .

Такой подход открывает прямой путь к рациональной организации функционирования восстанавливающего органа. Решая сформулированную задачу, мы будем существенно опираться на результаты работы [12], стремясь по возможности уточнить изложенные в ней идеи и выявить строгие границы их применимости.

Введём величину

$$Z = \sum_{i=1}^{n+1} a_i X_i, \tag{1.2.1}$$

рассматривая все x_i (за исключением $X_{n+1} = x_{n+1} \equiv -1$) из-за возможности появления ошибок в информационных каналах B_i в качестве случайных двоичных переменных X_i . Тогда, очевидно, и Z будет являться случайной величиной, принимающей значение z на вещественной оси, а $Y = \text{sgn } Z$ - случайной двоичной переменной.

Кроме того, составим случайную величину

$$\eta = X \cdot Z = \sum_{i=1}^{n+1} a_i (X \cdot X_i) = \sum_{i=1}^{n+1} \eta_i, \tag{1.2.2}$$

где X - поступившая на информационные каналы $B_i (i = \overline{1, n+1})$ двоичная переменная, принимающая значения $+1$ и -1 . Она трактуется в качестве случайной величины с реализациями x .

Реализации случайной величины η условимся обозначать через v , а реализации случайных величин η_i - через v_i . Легко заметить, что значения v дискретны и принадлежат вещественной оси, значения же v_i составляют либо $+a_i$, либо $-a_i$ для всех $i = \overline{1, n+1}$.

Вводя символы P, N и O соответственно для положительных, отрицательных и нулевых значений реализаций случайных величин Z и η , можно составить Таблицу 1.

Таблица 1. Доказательство тождества $X \text{sgn } Z \equiv \text{sgn}(XZ)$

X	Z	$\text{sgn}Z$	$X\text{sgn}Z$	XZ	$\text{sgn}(XZ)$
1	2	3	4	5	6
+1	P	+1	+1	P	+1
+1	O	O	O	O	O
+1	N	-1	-1	N	-1
-1	P	+1	-1	N	-1
-1	O	O	O	O	O
-1	N	-1	+1	P	+1

Из данных этой таблицы можно утверждать, что, если X есть случайная двоичная переменная, принимающая значения $+1$ и -1 , а функция $\text{sgn } Z$ определена согласно соотношению (1.2.1), то имеет место следующее тождество:

$$X \text{sgn } Z \equiv \text{sgn}(XZ).$$

На основании же последнего тождества вытекает справедливость цепочки преобразований:

$$\begin{aligned} Q = \text{Prob}\{Y \neq X\} &= \text{Prob}\{XY = -1\} = \text{Prob}\{X \text{sgn } Z = -1\} = \\ &= \text{Prob}\{\text{sgn}(XZ) = -1\} = \text{Prob}\{XZ < 0\} = \text{Prob}\{\eta < 0\}. \end{aligned}$$

Таким образом, вероятность Q того, что решение Y ошибочно, равна вероятности того, что случайная величина η отрицательна:

$$Q = \text{Prob}\{\eta < 0\}. \quad (1.2.3)$$

Следовательно, задача свелась к изучению распределения случайной величины (1.2.2). С этой целью обратимся, прежде всего, к случайным величинам $X \cdot X_i$ ($i = \overline{1, n+1}$). Легко видеть, что дискретная случайная величина $X \cdot X_i$ принимает значение -1 при $X_i = \neg X$ (где $\neg X$ означает инверсию двоичной переменной X) с вероятностью q_i и значение $+1$ при $X_i = X$ с вероятностью $1 - q_i$:

$$\left. \begin{aligned} \text{Prob}\{X \cdot X_i = -1\} &= \text{Prob}\{X_i \neq X\} = q_i \\ \text{Prob}\{X \cdot X_i = +1\} &= \text{Prob}\{X_i = X\} = 1 - q_i \end{aligned} \right\} \quad i = \overline{1, n+1} \quad (1.2.4)$$

В частности

$$\left. \begin{aligned} q_{n+1} &= \text{Prob}\{X \cdot X_{n+1} = -1\} = \text{Prob}\{X = +1\} \\ 1 - q_{n+1} &= \text{Prob}\{X \cdot X_{n+1} = +1\} = \text{Prob}\{X = -1\} \end{aligned} \right\} \quad (1.2.5)$$

так как $X_{n+1} = -1$. Из последних формул следует, что q_{n+1} есть априорная вероятность подачи на решающий орган для распознавания сигнала $X = +1$, т.е. q_{n+1} есть априорная вероятность появления $+1$ на выходе порогового элемента в качестве правильного сигнала. Аналогично, $1 - q_{n+1}$ есть априорная вероятность подачи на вход решающего органа сигнала $X = -1$, или, что то же самое, априорная вероятность появления -1 на выходе порогового элемента в качестве правильного решения.

Распределения $f_i(v_i)$ вероятностей дискретных случайных величин η_i задаются исходя из того, что $v_i = +a_i$ с вероятностью $1 - q_i$ и $v_i = -a_i$ с вероятностью q_i . Следовательно,

$$\left. \begin{aligned} f_i(v_i) &= q_i^{(a_i - v_i)/2a_i} \cdot (1 - q_i)^{(v_i + a_i)/2a_i} \\ &\quad v_i = +a_i, -a_i \\ &\quad i = \overline{1, n+1} \end{aligned} \right\} \quad (1.2.6)$$

Распределение $f(v)$ вероятностей дискретной случайной величины η , являющейся суммой независимых дискретных случайных величин η_i , найдётся как свёртка распределений $f_i(v_i)$ последних:

$$f(v) = \underset{i=1}{*}^{n+1} f_i(v_i), \quad (1.2.7)$$

где $*$ (логичнее было бы наложение знаков «+» сложения и « \times » умножения при наличии в нашем распоряжении подобного символа) – условное обозначение операции свёртки.

Совершенно очевидно, что - с учётом выражения (1.2.7) - искомая вероятность (1.2.3) ошибки на выходе решающего элемента примет вид:

$$Q = \sum_{v < 0} f(v) = \sum_{v < 0} \underset{i=1}{*}^{n+1} f_i(v_i), \quad (1.2.8)$$

где $f_i(v_i)$ определяется формулой (1.2.6), а суммирование ведётся по всем отрицательным дискретным значениям переменной v . Полное же число дискретных значений v составляет 2^{n+1} , поскольку

$$v = \tilde{a}_1 + \tilde{a}_2 + \dots + \tilde{a}_n + \tilde{a}_{n+1},$$

где выражение \tilde{a}_i равно либо $+a_i$, либо $-a_i$.

Согласно формуле (1.2.8), каждому дискретному значению v отвечает член Q_j ($j = \overline{1, 2^{n+1}}$), который представляет собой произведение определённых вероятностей:

$$Q_j \equiv f(v) = \left. \begin{aligned} & \equiv \prod_{i=1}^{n+1} f_i(v_i) \equiv \tilde{q}_1 \cdot \tilde{q}_2 \cdot \dots \cdot \tilde{q}_n \cdot \tilde{q}_{n+1} \\ & j = \overline{1, 2^{n+1}} \end{aligned} \right\}, \quad (1.2.9)$$

где

$$\tilde{q}_k = \begin{cases} q_k, & \text{если } v_k = -a_k \\ 1 - q_k, & \text{если } v_k = +a_k \end{cases}. \quad (1.2.10)$$

Следовательно,

$$Q = \sum_{v < 0} Q_j = \sum_{v < 0} \tilde{q}_1 \cdot \tilde{q}_2 \cdot \dots \cdot \tilde{q}_n \cdot \tilde{q}_{n+1}. \quad (1.2.11)$$

Для демонстрации этого алгоритма рассмотрим случай трёхвходового мажоритарного элемента, когда $n = 3$; $q_1 = q_2 = q_3 \equiv q$, $q_4 = 1/2$; $a_1 = a_2 = a_3 = 1$, $a_4 \equiv \Theta = 0$. Результаты расчётов вероятности ошибки Q трёхвходового мажоритарного элемента по алгоритму (1.2.11) представлены в Таблице 2.

Таблица 2. Расчёт вероятности ошибки Q

j	v_1	v_2	v_3	v_4	$v = v_1 + v_2 + v_3 + v_4$	$Q_j = \tilde{q}_1 \tilde{q}_2 \tilde{q}_3 \tilde{q}_4$
1	2	3	4	5	6	7
1	$-a_1=-1$	$-a_2=-1$	$-a_3=-1$	$-a_4=0$	-3	$q_1 q_2 q_3 q_4$
2	$-a_1=-1$	$-a_2=-1$	$-a_3=-1$	$+a_4=0$	-3	$q_1 q_2 q_3 (1 - q_4)$
3	$-a_1=-1$	$-a_2=-1$	$+a_3=+1$	$-a_4=0$	-1	$q_1 q_2 (1 - q_3) q_4$
4	$-a_1=-1$	$-a_2=-1$	$+a_3=+1$	$+a_4=0$	-1	$q_1 q_2 (1 - q_3)(1 - q_4)$
5	$-a_1=-1$	$+a_2=+1$	$-a_3=-1$	$-a_4=0$	-1	$q_1 (1 - q_2) q_3 q_4$
6	$-a_1=-1$	$+a_2=+1$	$-a_3=-1$	$+a_4=0$	-1	$q_1 (1 - q_2) q_3 (1 - q_4)$
7	$-a_1=-1$	$+a_2=+1$	$+a_3=+1$	$-a_4=0$	+1	
8	$-a_1=-1$	$+a_2=+1$	$+a_3=+1$	$+a_4=0$	+1	

9	$+a_1=+1$	$-a_2=-1$	$-a_3=-1$	$-a_4=0$	-1	$(1-q_1)q_2q_3q_4$
10	$+a_1=+1$	$-a_2=-1$	$-a_3=-1$	$+a_4=0$	-1	$(1-q_1)q_2q_3(1-q_4)$
11	$+a_1=+1$	$-a_2=-1$	$+a_3=+1$	$-a_4=0$	+1	
12	$+a_1=+1$	$-a_2=-1$	$+a_3=+1$	$+a_4=0$	+1	
13	$+a_1=+1$	$+a_2=+1$	$-a_3=-1$	$-a_4=0$	+1	
14	$+a_1=+1$	$+a_2=+1$	$-a_3=-1$	$+a_4=0$	+1	
15	$+a_1=+1$	$+a_2=+1$	$+a_3=+1$	$-a_4=0$	+3	
16	$+a_1=+1$	$+a_2=+1$	$+a_3=+1$	$+a_4=0$	+3	

С одной стороны известно, что для мажоритарного элемента

$$Q = \sum_{k=\lfloor n/2+1 \rfloor}^n C_n^k q^k (1-q)^{n-k}, \quad (1.2.12)$$

где C_n^k представляет собой число сочетаний из n элементов по k , причём $k = \lfloor n/2+1 \rfloor$ - является наибольшим целым числом, которое не превышает величины $\left(\frac{n}{2}+1\right)$. Такая зависимость, общеизвестно, трактуется как функция «антье» (entier – целая часть) аргумента $\left(\frac{n}{2}+1\right)$. Для рассмотренного примера $k = \lfloor n/2+1 \rfloor = 2$, причём

$$Q = 3q^2(1-q) + q^3. \quad (1.2.13)$$

С другой стороны этот результат можно сравнить с тем, что даёт описанный выше алгоритм (1.2.11). Необходимые для расчёта данные внесены в Таблицу 2.

Если сложить выражения Q_j из седьмого столбца, отвечающие отрицательным значениям v , то получим:

$$Q = (Q_1 + Q_2) + (Q_3 + Q_4 + Q_5 + Q_6) + (Q_9 + Q_{10}),$$

или

$$Q = q^3 + 2q^2(1-q) + q^2(1-q) = 3q^2(1-q) + q^3. \quad (1.2.14)$$

Таким образом, результаты (1.2.13) и (1.2.14) полностью совпадают.

Программа, которая реализует алгоритм (1.2.11) и позволяет точно рассчитать вероятность ошибочного восстановления двоичного сигнала пороговым элементом (искусственным, или формальным нейроном), на входе должна получать:

- число n входов порогового элемента;
- вероятность ошибки q_i ($i = \overline{1, n}$) на i -м входе;
- вес a_i ($i = \overline{1, n}$), приписываемый i -му входу;

- априорную вероятность q_{n+1} появления сигнала $+1$ в качестве правильного решения на выходе порогового элемента (т.е. априорную вероятность предъявления пороговому органу для распознавания, восстановления сигнала $+1$);

- величину $\Theta \equiv a_{n+1}$ порога.

Выходные же данные должны содержать:

- всю входную информацию;

- вероятность Q неправильного восстановления сигнала пороговым решающим органом с заданными значениями числа входов, вероятностей их ошибок, весов, порога и априорной вероятности предъявления к распознаванию сигнала $+1$.

2. Минимальная верхняя оценка вероятности ошибки восстановления

Одновременно для вероятности Q полезно иметь верхнюю оценку в замкнутой аналитической форме. Вывод такой оценки, полученной в работе [12], страдает существенными изъянами. Ниже предлагается новый подход к выводу такой оценки.

Известно [102], что производящая функция $\gamma_v(S)$ факториальных моментов суммы η независимых случайных величин η_i равна произведению производящих функций $\gamma_{v_i}(S)$ факториальных моментов отдельных слагаемых:

$$\gamma_v(S) = \prod_{i=1}^{n+1} \gamma_{v_i}(S), \quad (2.1)$$

где

$$\gamma_v(S) = M[S^\eta] = \sum_v S^v \cdot f(v), \quad (2.2)$$

$$\left. \begin{aligned} \gamma_{v_i}(S) &= M[S^{\eta_i}] = \sum_v S^{v_i} \cdot f_i(v_i) \\ i &= \overline{1, n+1} \end{aligned} \right\}. \quad (2.3)$$

Здесь M - символ математического ожидания, а S - любое комплексное число, для которого ряды (2.2) и (2.3) существуют в смысле абсолютной сходимости в некотором интервале действительной оси, включающем точку $S = 1$.

Поскольку в соотношении (2.3) суммирование ведётся по двум значениям $+a_i$ и $-a_i$ величины v_i , с учётом выражений (1.2.6) получим:

$$\left. \begin{aligned} \gamma_{v_i}(S) &= (1 - q_i)S^{a_i} + q_i S^{-a_i} \\ i &= \overline{1, n+1} \end{aligned} \right\}. \quad (2.4)$$

Подставляя же (2.4) в (2.1), будем иметь:

$$\gamma_v(S) = \prod_{i=1}^{n+1} [(1 - q_i)S^{a_i} + q_i S^{-a_i}].$$

При $v < 0$ величина S^v удовлетворяет условию

$$S^v = \frac{1}{S^{|v|}} > 1,$$

если только

$$0 < S < 1. \quad (2.5)$$

Допустим, что неравенство (2.5) выполнено. Тогда справедливо следующее соотношение:

$$Q = \sum_{v<0} f(v) < \sum_{v<0} S^v \cdot f(v).$$

Поскольку слагаемые $S^v \cdot f(v)$ неотрицательны, имеет место неравенство

$$\sum_{v<0} S^v \cdot f(v) \leq \sum_v S^v \cdot f(v).$$

Следовательно,

$$Q < \gamma_v(S). \quad (2.6)$$

Правую часть этого выражения и можно принять за верхнюю оценку Q^+ вероятности ошибки Q порогового элемента:

$$Q^+ = \prod_{i=1}^{n+1} [(1-q_i)S^{a_i} + q_i S^{-a_i}].$$

Последнее соотношение представимо и в следующей эквивалентной форме:

$$Q^+ = \prod_{i=1}^{n+1} Q_i^+ = \prod_{i=1}^{n+1} \left[(1-q_i)w_i + \frac{q_i}{w_i} \right], \quad (2.7)$$

где

$$Q_i^+ = (1-q_i)w_i + \frac{q_i}{w_i},$$

причём

$$w_i = S^{a_i}, \quad 0 < w_i < \infty, \quad (i = \overline{1, n+1}). \quad (2.8)$$

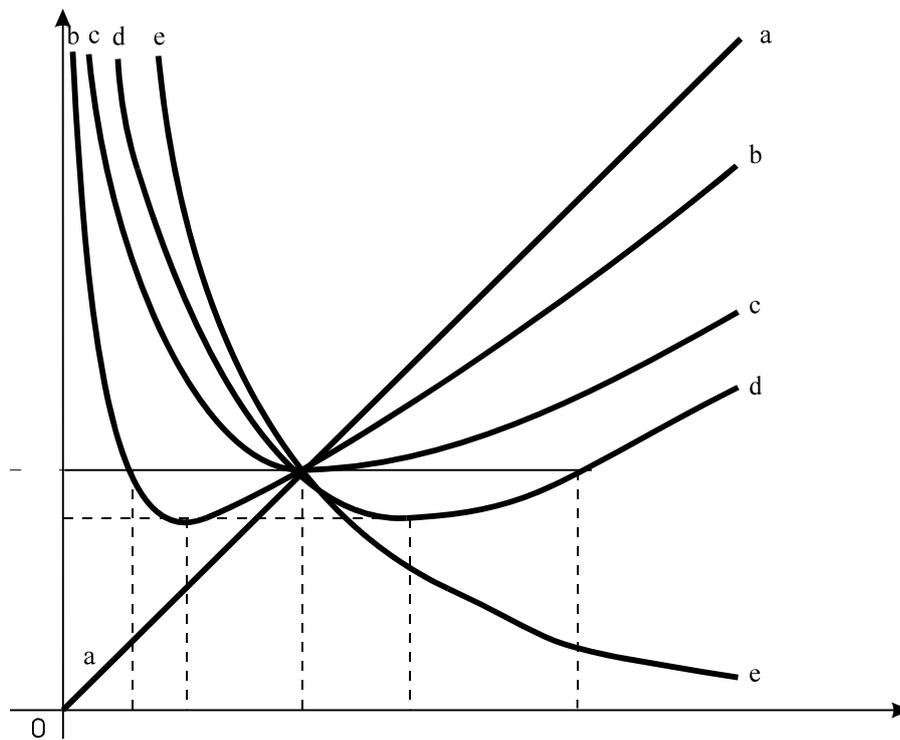
Графики зависимости Q_i^+ от w_i для ряда значений вероятности q_i приведены на рисунке 5. В частности здесь прямая a отвечает значению $q_i = 0$, кривая b - значению $0 < q_i < 1/2$, кривая c - значению $q_i = 1/2$, кривая d - значению $1/2 < q_i < 1$, наконец, кривая e - значению $q_i = 1$.

Пунктирная горизонталь отвечает значению

$$Q_i^+ = Q_{\min i}^+ = 2 \cdot \sqrt{q_i(1-q_i)}.$$

Сплошная горизонталь отвечает значению

$$Q_i^+ = 1.$$


 Рис. 5 Графическое представление зависимости Q_i^+ от w_i

На оси абсцисс пунктирные вертикали отмечают слева направо характерные значения w_i :

$$Q_i^+ = 1, \text{ если } w_i = \frac{q_i}{1-q_i} < 1;$$

$$Q_i^+ = Q_{\min i}^+, \text{ если } w_i = \sqrt{\frac{q_i}{1-q_i}} < 1;$$

$$Q_i^+ = 1, \text{ если } w_i = 1;$$

$$Q_i^+ = Q_{\min i}^+, \text{ если } w_i = \sqrt{\frac{q_i}{1-q_i}} > 1;$$

$$Q_i^+ = 1, \text{ если } w_i = \frac{q_i}{1-q_i} > 1.$$

Естественно попытаться найти минимум Q_{\min}^+ выражения (2.7) и значения w_{0i} величин w_i , доставляющие верхней оценке вероятности ошибки порогового элемента минимум. Для этого воспользуемся условиями

$$\left. \begin{array}{l} \frac{\partial Q^+}{\partial w_i} = 0 \\ i = \overline{1, n+1} \end{array} \right\},$$

которые дают:

$$w_{0i} = \sqrt{\frac{q_i}{1-q_i}}, \quad (i = \overline{1, n+1}). \quad (2.9)$$

Подставляя (2.9) в (2.7), получим формулу для минимальной верхней оценки вероятности ошибки порогового элемента:

$$Q_{\min}^+ = 2^{n+1} \cdot \prod_{i=1}^{n+1} [\sqrt{q_i(1-q_i)}],$$

или, что то же самое,

$$Q_{\min}^+ = \exp\left(-\sum_{i=1}^{n+1} A(q_i)\right), \quad (2.10)$$

где

$$A(q_i) = \left| \ln\left[2\sqrt{q_i(1-q_i)}\right] \right|.$$

Из формулы (2.10) с учётом неотрицательности величин $A(q_i)$ видно, что увеличение числа входов n порогового элемента, если только вероятности ошибок на этих входах не равны $1/2$, ведёт к монотонному уменьшению по экспоненте даже минимальной верхней оценки вероятности неправильного распознавания.

В этом результате обнаруживается глубокая внутренняя связь с теоремой Шеннона [103], которая утверждает, что число сообщений заданной длины n (длительности τ), составленных из отдельных символов – как в отсутствие так и при наличии фиксированных и вероятностных ограничений (в последнем случае при условии эргодичности источника) – с увеличением n (или τ) возрастает по асимптотически показательному закону. Эта связь видится нам в следующем: по какому закону с увеличением n резервированных двоичных каналов возрастает на входах решающего элемента количество той информации, на основе которой может приниматься решение y , по такому же (степенному) закону убывает при этом вероятность того, что принятое решение ошибочно.

Оптимальные значения весов a_i ($i = \overline{1, n+1}$) находят из соотношений (2.9) с учётом обозначений (2.8):

$$a_i = \frac{1}{2 \ln S} \cdot \ln \frac{q_i}{1-q_i} \left. \vphantom{a_i} \right\}_{i = \overline{1, n+1}}.$$

Поскольку S удовлетворяет условию (2.5), величина $\ln S < 0$. Поэтому

$$a_i = k \cdot \ln \frac{1-q_i}{q_i} \left. \vphantom{a_i} \right\}_{i = \overline{1, n+1}}, \quad (2.11)$$

где

$$k = \frac{1}{2|\ln S|} \left. \vphantom{k} \right\}_{0 < k < \infty}. \quad (2.12)$$

Следовательно, оптимальные веса a_i ($i = \overline{1, n+1}$), обеспечивающие минимальную вероятность ошибки порогового органа, определены с точностью до общего положительного множителя k .

Характер зависимости весов a_i , а также оценок $Q_{\min i}^+$ и $A(q_i)$ от q_i представлен на рисунке 6, причём

$$Q_{\min i}^+ = 2\sqrt{q_i(1-q_i)}$$

и для значений $q_i > 0.8$ выражение $2\sqrt{q_i(1-q_i)}$ меньше q_i .

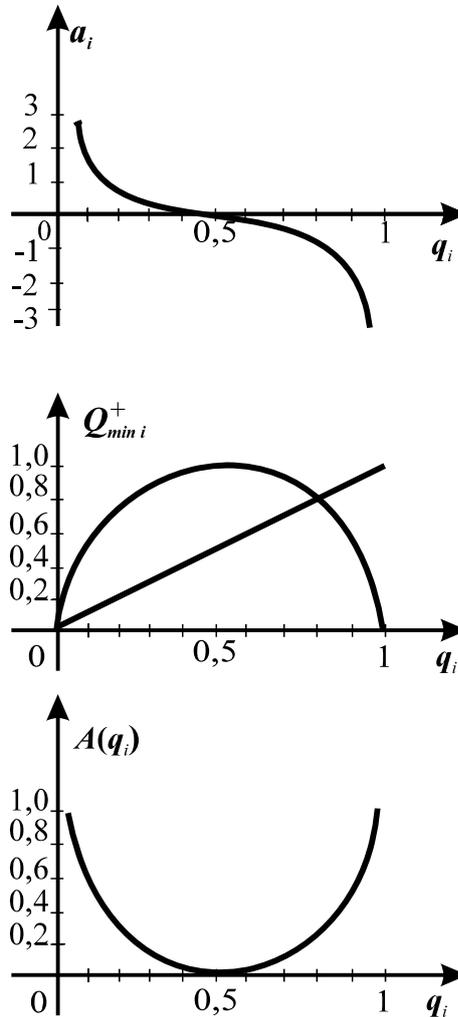


Рис. 6 Зависимость весов a_i и оценок $Q_{\min i}^+$, $A(q_i)$ от вероятности ошибки q_i

Применение полученных результатов удобнее всего проиллюстрировать примером.

Предположим, что вероятности ошибок на $n = 8$ входах решающего элемента в начале эксплуатации распределены следующим образом:

$$q_1 = q_2 = 10^{-5}, \quad q_3 = q_4 = 10^{-3}, \quad q_5 = q_6 = 10^{-1}, \quad q_7 = q_8 = 6 \cdot 10^{-1}.$$

К концу же срока службы численные значения этих характеристик составляют:

$$q_1 = 10^{-3}, \quad q_2 = 10^{-1}, \quad q_3 = q_4 = 5 \cdot 10^{-1}, \quad q_5 = q_6 = q_7 = q_8 = 8 \cdot 10^{-1}.$$

Априорная вероятность q_0 предъявления к распознаванию сигнала +1 на протяжении всего времени постоянна и составляет $1/2$.

Полагается, что веса и порог выбираются оптимально, в соответствии с текущими значениями вероятностей $q_i (i = \overline{1,9})$.

Точное значение вероятности ошибки порогового элемента в начале эксплуатации составляет $Q = 3.397 \cdot 10^{-9}$, а верхняя минимальная оценка равна $Q_{\min}^+ = 5.524 \cdot 10^{-8}$. К концу же срока службы эти величины соответственно равны $7.912 \cdot 10^{-4}$ и $1.554 \cdot 10^{-2}$.

Приведённые данные свидетельствуют о том, что в аналитических расчётах использование оценки (2.10) требует известной осторожности, так как она может оказаться слишком грубой (а именно, завышенной).

3. Асимптотическая оценка вероятности ошибки при мажоритарном резервировании

Рассмотрим решающий орган мажоритарного типа, в котором вероятности ошибок информационных каналов одинаковы и равны $q_i \equiv q = 1 - p$, веса $a_i = 1 (i = \overline{1, n})$, а порог $a_{n+1} \equiv \Theta = 0$. Обозначим через ξ число ошибок на n входах решающего элемента. Если эти ошибки независимы, распределение случайной величины ξ будет подчиняться биномиальному закону, т.е.

$$\text{Prob}\{\xi = k\} = C_n^k q^k (1 - q)^{n-k} \quad \left. \vphantom{\text{Prob}\{\xi = k\}} \right\} \quad k = 0, 1, 2, \dots, n \quad (3.1)$$

В этих предположениях вероятность Q ошибки на выходе мажоритарного элемента равна вероятности того, что случайная величина ξ окажется не меньше половины общего числа его входов, иначе говоря, не меньше, чем $\left\lfloor \frac{n}{2} + 1 \right\rfloor$, где $\left\lfloor \frac{n}{2} + 1 \right\rfloor$ - наибольшая целая часть величины $\frac{n}{2} + 1$:

$$Q = \text{Prob}\left\{\xi \geq \left\lfloor \frac{n}{2} + 1 \right\rfloor\right\}. \quad (3.2)$$

Следовательно,

$$Q = \sum_{k=\left\lfloor \frac{n}{2} + 1 \right\rfloor}^n C_n^k q^k (1 - q)^{n-k}. \quad (3.3)$$

Выражение (3.3) является точным, однако оно не даёт достаточно наглядного представления о характере зависимости

$$Q = f(q, n).$$

Поэтому попытаемся получить для верхней границы вероятности (3.3) асимптотическую оценку, отвечающую условию $n \rightarrow \infty$.

Легко видеть, что в правой части выражения (3.3) отношение последующего члена суммы к предыдущему не остаётся постоянным, поскольку оно зависит от порядкового номера j , т.е.

$$\frac{n-j}{j+1} \cdot \frac{q}{1-q} \equiv \gamma_j,$$

где

$$j = \left\lfloor \frac{n}{2} + 1 \right\rfloor, \left\lfloor \frac{n}{2} + 1 \right\rfloor + 1, \dots, n - 1.$$

Следовательно,

$$\max_j \{ \gamma_j \} = \gamma_{\left\lfloor \frac{n}{2} + 1 \right\rfloor},$$

причём

$$\gamma_{\left\lfloor \frac{n}{2} + 1 \right\rfloor} = \frac{n - \left\lfloor \frac{n}{2} + 1 \right\rfloor}{\left\lfloor \frac{n}{2} + 1 \right\rfloor + 1} \cdot \frac{q}{1 - q}. \quad (3.4)$$

Легко доказать, что

$$0 < \gamma_{\left\lfloor \frac{n}{2} + 1 \right\rfloor} < 1, \quad (3.5)$$

если

$$0 < q < \frac{\left\lfloor \frac{n}{2} + 1 \right\rfloor + 1}{n + 1} \equiv q_0(n), \quad (3.6)$$

где, как нетрудно видеть, для конечных значений n

$$\frac{1}{2} < q_0(n) \leq 1.$$

Однако при $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} q_0(n) = \frac{1}{2},$$

и условие (3.6) справедливости соотношения (3.5) приобретает следующий вид:

$$0 < q < \frac{1}{2}. \quad (3.7)$$

Впредь будем предполагать, что неравенство (3.7) выполнено, и рассмотрим бесконечно убывающую геометрическую прогрессию со знаменателем $\gamma_{\left\lfloor \frac{n}{2} + 1 \right\rfloor}$ и первым членом, равным величине

$$b_0 = C_n^{\left\lfloor \frac{n}{2} + 1 \right\rfloor} \cdot q^{\left\lfloor \frac{n}{2} + 1 \right\rfloor} \cdot (1 - q)^{n - \left\lfloor \frac{n}{2} + 1 \right\rfloor}.$$

Ясно, что вероятность Q не может превосходить суммы членов этой прогрессии, т.е.

$$Q \leq \frac{1}{1 - \frac{n - \left\lfloor \frac{n}{2} + 1 \right\rfloor}{\left\lfloor \frac{n}{2} + 1 \right\rfloor + 1} \cdot \frac{q}{1 - q}} \cdot C_n^{\left\lfloor \frac{n}{2} + 1 \right\rfloor} \cdot q^{\left\lfloor \frac{n}{2} + 1 \right\rfloor} \cdot (1 - q)^{n - \left\lfloor \frac{n}{2} + 1 \right\rfloor}. \quad (3.8)$$

Логарифмируя обе части последнего неравенства, получим:

$$\ln Q \leq \ln \frac{1}{1 - \frac{\binom{n-1}{\frac{n}{2}+1} \cdot q}{\binom{n}{\frac{n}{2}+1} \cdot (1-q)}} + \ln C_n^{\lfloor \frac{n}{2} \rfloor + 1} + \left\lfloor \frac{n}{2} \right\rfloor + 1 \left[\ln q + \left(n - \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) \ln(1-q) \right]. \quad (3.9)$$

Дальнейший анализ этого соотношения проведём для случаев чётного и нечётного n раздельно.

Пусть n чётно. Следовательно, $\left\lfloor \frac{n}{2} \right\rfloor + 1 = \frac{n}{2} + 1$ и, таким образом,

$$\ln Q \leq \ln \frac{1}{1 - \frac{\frac{n-1}{2} \cdot q}{\frac{n}{2} + 2} \cdot \frac{q}{1-q}} + \ln C_n^{\frac{n}{2}+1} + \left(\frac{n}{2} + 1 \right) \ln q + \left(\frac{n}{2} - 1 \right) \ln(1-q).$$

Разделив обе части данного неравенства на n и перейдя к пределу при $n \rightarrow \infty$, будем иметь:

$$\lim_{n \rightarrow \infty} \frac{\ln Q}{n} \leq \ln \sqrt{q(1-q)} + \lim_{n \rightarrow \infty} \frac{1}{n} \ln C_n^{\frac{n}{2}+1}, \quad (3.10)$$

где

$$C_n^{\frac{n}{2}+1} = \frac{n!}{\left(\frac{n}{2} + 1\right)! \left(\frac{n}{2} - 1\right)!}.$$

Следовательно,

$$\ln C_n^{\frac{n}{2}+1} = \ln n! - \ln \left(\frac{n}{2} + 1\right)! - \ln \left(\frac{n}{2} - 1\right)!.$$

По формуле Стирлинга

$$\begin{aligned} \ln n! &= \left(n + \frac{1}{2}\right) \ln n - n + \ln \sqrt{2\pi}, \\ \ln \left(\frac{n}{2} + 1\right)! &= \left(\frac{n}{2} + \frac{3}{2}\right) \ln \left(\frac{n}{2} + 1\right) - \left(\frac{n}{2} + 1\right) + \ln \sqrt{2\pi}, \\ \ln \left(\frac{n}{2} - 1\right)! &= \left(\frac{n}{2} - \frac{1}{2}\right) \ln \left(\frac{n}{2} - 1\right) - \left(\frac{n}{2} - 1\right) + \ln \sqrt{2\pi}. \end{aligned}$$

Поэтому

$$\ln C_n^{\frac{n}{2}+1} = \left(n + \frac{1}{2}\right) \ln n - \left(\frac{n}{2} + \frac{3}{2}\right) \ln \left(\frac{n}{2} + 1\right) + \left(\frac{n}{2} - \frac{1}{2}\right) \ln \left(\frac{n}{2} - 1\right) - \ln \sqrt{2\pi}.$$

Разделив обе части последнего равенства на n и устремив эту величину к бесконечности, получим:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln C_n^{\frac{n}{2}+1} = \lim_{n \rightarrow \infty} \left[\ln n - \frac{1}{2} \ln \left(\frac{n}{2} + 1\right) - \frac{1}{2} \ln \left(\frac{n}{2} - 1\right) \right],$$

откуда

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln C_n^{\frac{n}{2}+1} = \ln 2. \quad (3.11)$$

С учётом этого соотношения неравенство (3.10) окончательно примет следующий вид:

$$\lim_{n \rightarrow \infty} \frac{\ln Q}{n} \leq \ln 2 \sqrt{q(1-q)}. \quad (3.12)$$

Пусть сейчас n нечётно. Следовательно, $\left\lceil \frac{n}{2} + 1 \right\rceil = \frac{n}{2} + \frac{1}{2}$ и, таким образом,

$$\ln Q \leq \ln \frac{1}{1 - \frac{\frac{n}{2} - 1}{\frac{n}{2} + \frac{3}{2}} \cdot \frac{q}{1-q}} + \ln C_n^{\frac{n}{2} + \frac{1}{2}} + \left(\frac{n}{2} + \frac{1}{2}\right) \cdot \ln q + \left(\frac{n}{2} - \frac{1}{2}\right) \cdot \ln(1-q).$$

Поступая далее совершенно аналогично предыдущему случаю, получим:

$$\lim_{n \rightarrow \infty} \frac{\ln Q}{n} \leq \ln \sqrt{q(1-q)} + \lim_{n \rightarrow \infty} \frac{1}{n} \ln C_n^{\frac{n}{2}+1},$$

где, как нетрудно показать, по-прежнему

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln C_n^{\frac{n}{2} + \frac{1}{2}} = \ln 2.$$

Следовательно, для нечётных n , равно как и для чётных,

$$\lim_{n \rightarrow \infty} \frac{\ln Q}{n} \leq \ln 2 \sqrt{q(1-q)}.$$

На основании вышеизложенного заключаем, что для верхней асимптотической границы Q_a^+ вероятности Q имеет место следующее соотношение:

$$\frac{\ln Q_a^+}{n} = \ln 2 \sqrt{q(1-q)}. \quad (3.13)$$

Для выполнения условия

$$\ln 2 \sqrt{q(1-q)} < 0$$

необходимо и достаточно, чтобы выполнялось неравенство

$$2\sqrt{q(1-q)} < 1.$$

Легко показать, что последнему неравенству удовлетворяют все значения q из области $0 \leq q \leq 1$, за исключением точки $q = 1/2$, в которой $\ln 2 \sqrt{q(1-q)} = 0$.

Таким образом,

$$\left. \begin{array}{l} \ln 2 \sqrt{q(1-q)} \leq 0 \\ 0 \leq q \leq 1 \end{array} \right\}.$$

Вводя обозначение

$$A(q) = \left| \ln 2\sqrt{q(1-q)} \right|, \tag{3.14}$$

соотношению (3.13) окончательно можно придать следующий вид:

$$\frac{\ln Q_a^+}{n} = -A(q),$$

откуда

$$Q_a^+ = e^{-A(q) \cdot n}, \tag{3.15}$$

либо

$$\ln Q_a^+ = -A(q) \cdot n. \tag{3.16}$$

Таким образом, для всех значений q , удовлетворяющих условию (3.7), при котором величина $A(q)$ строго положительна, асимптотическая оценка (3.15) для верхней границы вероятности ошибки на выходе мажоритарного элемента с ростом n убывает по экспоненциальному закону.

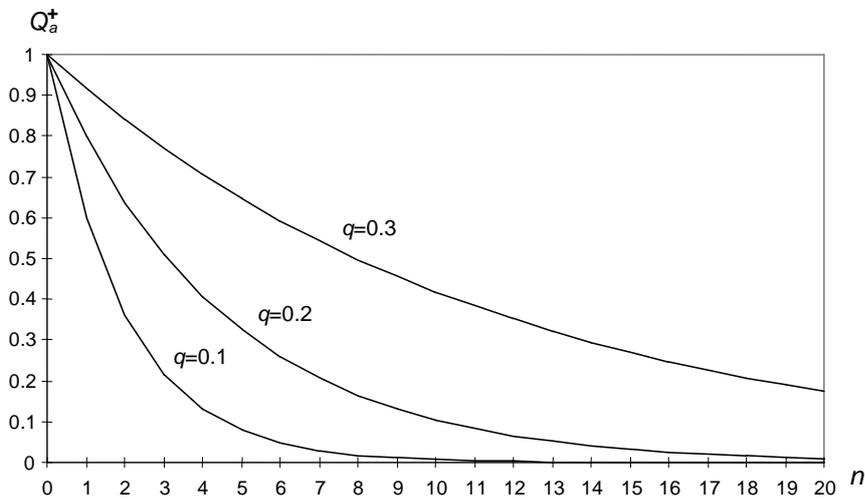


Рис. 7 Зависимость вероятности Q_a^+ от числа входов n при некоторых значениях q ($0 < q < 1/2$)

На рисунках 7 и 8 даётся графическое представление зависимостей (3.15) и (3.16) соответственно для ряда значений q , удовлетворяющих условию $0 < q < 1/2$.

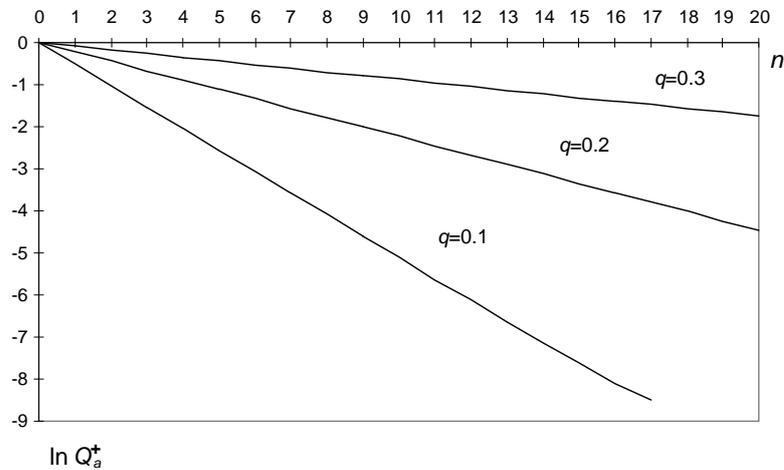


Рис.8 Зависимость натурального логарифма вероятности Q_a^+ от числа входов n при некоторых значениях q ($0 < q < 1/2$)

В следующих разделах мы более подробно остановимся на вопросе определения оптимальных входных весов порогового решающего элемента на основе метода обобщенного (махаланобисова) расстояния, а также байесовского и энтропийного подходов.

4. Оптимизация весов: метод обобщенного (махаланобисова) расстояния

Приступая к решению задачи оптимизации весов, будем трактовать её как задачу классификации, т.е. отнесения входного сигнала x к одному из двух классов Ω_1 и Ω_2 на основе n версий x_1, x_2, \dots, x_n этого сигнала.

При таком подходе восстанавливаемый сигнал следует рассматривать в качестве случайной величины X с реализацией x . Характеристиками этой величины служат n случайных величин X_1, X_2, \dots, X_n . Полезно их рассматривать как случайный вектор, т.е. упорядоченный набор из n чисел, расположенных в виде столбца:

$$\bar{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = (X_1, X_2, \dots, X_n)',$$

где знак «'» означает транспонирование. Следовательно, каждый элемент X_i ($i = \overline{1, n}$) служит компонентой случайного вектора.

Реализация случайного вектора \bar{X} записывается в виде вектора наблюдений

$$\bar{x} = (x_1, x_2, \dots, x_n)',$$

где компоненты вектора \bar{x} суть реализации x_1, x_2, \dots, x_n случайных величин X_1, X_2, \dots, X_n соответственно.

Вектор \bar{X} с компонентами X_1, X_2, \dots, X_n описывается совместной функцией распределения

$$f(\bar{x}) = \text{Prob}\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\},$$

где возможные значения x_i ($i = \overline{1, n}$) составляют -1 и $+1$.

Легко видеть, что в классе Ω_1 вектор \bar{X} имеет распределение

$$f_1(\bar{x}) = \prod_{i=1}^n q_i^{\frac{1-x_i}{2}} \cdot (1-q_i)^{\frac{x_i+1}{2}}, \quad (4.1)$$

а в классе Ω_2 - распределение

$$f_2(\bar{x}) = \prod_{i=1}^n q_i^{\frac{x_i+1}{2}} \cdot (1-q_i)^{\frac{1-x_i}{2}}. \quad (4.2)$$

Эти соотношения справедливы в предположении, что компоненты X_1, X_2, \dots, X_n вектора \bar{X} независимы.

Математическое ожидание $\mu_i = M[X_i]$ каждой из компонент X_i получают из частных распределений величин X_i . Эти n математических ожиданий можно представить в виде вектора $\vec{\mu}$ средних значений:

$$\vec{\mu} = M[\vec{X}] = (\mu_1, \mu_2, \dots, \mu_n)'.$$

В частности в классе Ω_1 центр распределения вероятностей вектора \vec{X} задаётся вектором

$$\vec{\mu}_1 = (\mu_{11}, \mu_{12}, \dots, \mu_{1n})', \tag{4.3}$$

а в классе Ω_2 - вектором

$$\vec{\mu}_2 = (\mu_{21}, \mu_{22}, \dots, \mu_{2n})', \tag{4.4}$$

где

$$\left. \begin{aligned} \mu_{1i} &= 1 - 2q_i \\ \mu_{2i} &= 2q_i - 1 \\ i &= \overline{1, n} \end{aligned} \right\}. \tag{4.5}$$

Из частных распределений для X_i рассчитывают дисперсии σ_i^2 случайных величин X_i ($i = \overline{1, n}$), а из совместного распределения компонент X_i и X_j вычисляют ковариацию σ_{ij} величин X_i и X_j :

$$\sigma_{ij} = M\left[(X_i - \mu_i)(X_j - \mu_j)\right] = \begin{cases} \sigma_i^2, & \text{когда } i = j \\ \sigma_{ij}, & \text{когда } i \neq j \end{cases}. \tag{4.6}$$

Заметим, что $\sigma_{ii} = \sigma_i^2$ и $\sigma_{ij} = \sigma_{ji}$. Дисперсии и ковариации образуют в совокупности ковариационную матрицу, являющуюся обобщением понятия дисперсии одномерной случайной величины:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \dots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \dots & \sigma_n^2 \end{bmatrix}.$$

Поскольку в рассматриваемом случае компоненты X_1, X_2, \dots, X_n вектора \vec{X} взаимно независимы, то $\sigma_{ij} = 0$ для всех $i \neq j$ и Σ представляет собой диагональную матрицу:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 \dots 0 \\ 0 & \sigma_2^2 & 0 \dots 0 \\ \vdots & \vdots & \vdots \dots \vdots \\ 0 & 0 & 0 \dots \sigma_n^2 \end{bmatrix}.$$

Легко показать, что в обоих классах дисперсии σ_i^2 одинаковы и определяются соотношениями

$$\left. \begin{aligned} \sigma_{(1)i}^2 = \sigma_{(2)i}^2 = \sigma_i^2 = 4q_i(1-q_i) \\ i = \overline{1, n} \end{aligned} \right\}. \quad (4.7)$$

Следовательно, совпадают и ковариационные матрицы:

$$\Sigma_1 = \Sigma_2 \equiv \Sigma. \quad (4.8)$$

Таким образом, при изучении порогового элемента параметры $\bar{\mu}_1$, $\bar{\mu}_2$ и Σ можно считать заданными. Из принципа же его функционирования следует, что в пороговом органе вычисляют линейную комбинацию наблюдений

$$z = a_1x_1 + a_2x_2 + \dots + a_nx_n,$$

называемую линейной дискриминантной функцией. Вектор наблюдений \vec{x} относят к классу Ω_1 , если $z > \Theta$, и к классу Ω_2 , если $z < \Theta$. При $z = 0$ решение не выносят.

Введём в рассмотрение случайную величину Z по соотношению

$$Z = \sum_{i=1}^n a_i X_i, \quad (4.9)$$

или

$$Z_i = \sum_{i=1}^n Z_i, \quad (4.10)$$

где

$$Z_i = a_i X_i. \quad (4.11)$$

Если наблюдение \vec{x} поступило из Ω_1 , то величина суммы (4.9) имеет распределение

$$F_1(z) = \underset{i=1}{*}^n f_{1i}(z_i), \quad (4.12)$$

где * - знак операции свёртки, а

$$f_{1i}(z_i) = q_i^{\frac{a_i - z_i}{2a_i}} (1 - q_i)^{\frac{z_i + a_i}{2a_i}}, \quad (2.13)$$

причём z_i равны либо $+a_i$, либо $-a_i$.

Математическое ожидание случайной величины Z составляет

$$m_1 = \sum_{i=1}^n a_i \mu_{1i} = \sum_{i=1}^n a_i (1 - 2q_i). \quad (4.14)$$

Аналогично, для наблюдения $\vec{x} = (x_1, x_2, \dots, x_n)$ из класса Ω_2 величина Z имеет распределение вероятностей

$$F_2(z) = \underset{i=1}{*}^n f_{2i}(z_i), \quad (4.15)$$

где

$$f_{2i}(z_i) = q_i^{\frac{z_i + a_i}{2a_i}} (1 - q_i)^{\frac{a_i - z_i}{2a_i}} \quad (4.16)$$

и возможные значения z_i составляют $+a_i$ и $-a_i$.

Если наблюдение \bar{x} поступило из класса Ω_2 , математическое ожидание случайной величины Z определяется формулой

$$m_2 = \sum_{i=1}^n a_i \mu_{2i} = \sum_{i=1}^n a_i (2q_i - 1). \quad (4.17)$$

Сравнивая соотношения (4.14) и (4.17), легко заключить, что

$$m_1 = -m_2. \quad (4.18)$$

Анализируя же формулы (4.13) и (4.16), находим:

$$\left. \begin{aligned} f_{2i}(z_i) &= f_{1i}(-z_i) \\ i &= \overline{1, n} \end{aligned} \right\}. \quad (4.19)$$

Следовательно,

$$F_2(z) = F_1(-z). \quad (4.20)$$

Дисперсия σ_z^2 случайной величины Z в обоих случаях одинакова и определяется формулой

$$\sigma_z^2 = \sum_{i=1}^n \sum_{j=1}^n a_i \sigma_{ij} a_j, \quad (4.21)$$

что с учётом соотношений (4.7), при $\sigma_{ij} = 0$, даёт:

$$\sigma_z^2 = \sum_{i=1}^n 4a_i^2 q_i (1 - q_i). \quad (4.22)$$

Эвристические соображения диктуют выбор таких весов a_1, a_2, \dots, a_n , при которых математические ожидания m_1 и m_2 как можно дальше отстоят друг от друга, а дисперсия σ_z^2 минимальна.

Для этого достаточно в качестве целевой функции ввести обобщённое расстояние Махаланобиса [104]:

$$\rho = \frac{(m_1 - m_2)^2}{\sigma_z^2}, \quad (4.23)$$

которое в условиях рассматриваемой нами задачи имеет вид:

$$\rho = \frac{\left[\sum_{i=1}^n a_i (\mu_{1i} - \mu_{2i}) \right]^2}{\sum_{i=1}^n \sum_{j=1}^n a_i \sigma_{ij} a_j}. \quad (4.24)$$

Веса $a_i (i = \overline{1, n})$, доставляющие этому выражению максимум, удовлетворяют следующей системе уравнений:

$$\left. \begin{aligned} \frac{\partial \rho}{\partial a_i} &= 0 \\ i &= \overline{1, n} \end{aligned} \right\}. \quad (4.25)$$

С учётом выражения (4.24) получим:

$$\frac{\mu_{1s} - \mu_{2s}}{\sum_{j=1}^n \sigma_{sj} a_j} = \frac{\sum_{i=1}^n a_i (\mu_{1i} - \mu_{2i})}{\sum_{i=1}^n a_i \sum_{j=1}^n \sigma_{ij} a_j}. \quad (4.26)$$

Всякий вектор $(a_1, a_2, \dots, a_n)'$, который удовлетворяет системе уравнений

$$\left. \begin{aligned} \sum_{j=1}^n \sigma_{sj} a_j &= k(\mu_{1s} - \mu_{2s}) \\ s &= \overline{1, n} \end{aligned} \right\}, \quad (1.27)$$

где k – произвольная постоянная, будет решением и системы уравнений (4.26).

Учитывая же в системе (4.23), что $\sigma_{ij} = 0$ при $i \neq j$ и $\sigma_{ii} = \sigma_i^2$, получим:

$$\left. \begin{aligned} a_s \sigma_s^2 &= k(\mu_{1s} - \mu_{2s}) \\ s &= \overline{1, n} \end{aligned} \right\},$$

откуда

$$\left. \begin{aligned} a_s &= k \frac{\mu_{1s} - \mu_{2s}}{\sigma_s^2} \\ s &= \overline{1, n} \end{aligned} \right\}. \quad (4.28)$$

С учётом выражений (4.5) и (4.7) окончательно будем иметь:

$$\left. \begin{aligned} a_i &= k \frac{1 - 2q_i}{2q_i(1 - q_i)} \\ i &= \overline{1, n} \end{aligned} \right\}. \quad (4.29)$$

Если при $q_i < 1/2$ желательно иметь положительный вес, а при $q_i > 1/2$ – отрицательный, то постоянная k должна удовлетворять условию $0 < k < \infty$.

При этих весах обобщённое расстояние ρ равно разности (точнее, абсолютной величине разности) математических ожиданий m_1 и m_2 суммы Z для классов Ω_1 и Ω_2 :

$$\rho = |m_1 - m_2| = \sum_{i=1}^n \frac{(1 - 2q_i)^2}{q_i(1 - q_i)}. \quad (4.30)$$

Следовательно, когда веса a_i выбраны согласно соотношениям (4.29), имеет место равенство

$$\frac{\sigma_z^2}{|m_1 - m_2|} = 1. \quad (4.31)$$

Из формулы (4.30) видно, что увеличение числа входов n порогового элемента, если только вероятности ошибок на этих входах не равны $1/2$, ведёт к монотонному росту обобщённого расстояния ρ и, следовательно, к уменьшению вероятности неправильного восстановления сигнала решающим органом. В частности в мажоритарном элементе с $q_1 = q_2 = \dots = q_n \equiv q$ для обобщённого расстояния ρ имеем:

$$\rho = \frac{(1-2q)^2}{q(1-q)} \cdot n. \tag{4.32}$$

Семейство зависимостей $\rho(q)$ для ряда значений n приведено на рисунке 9.

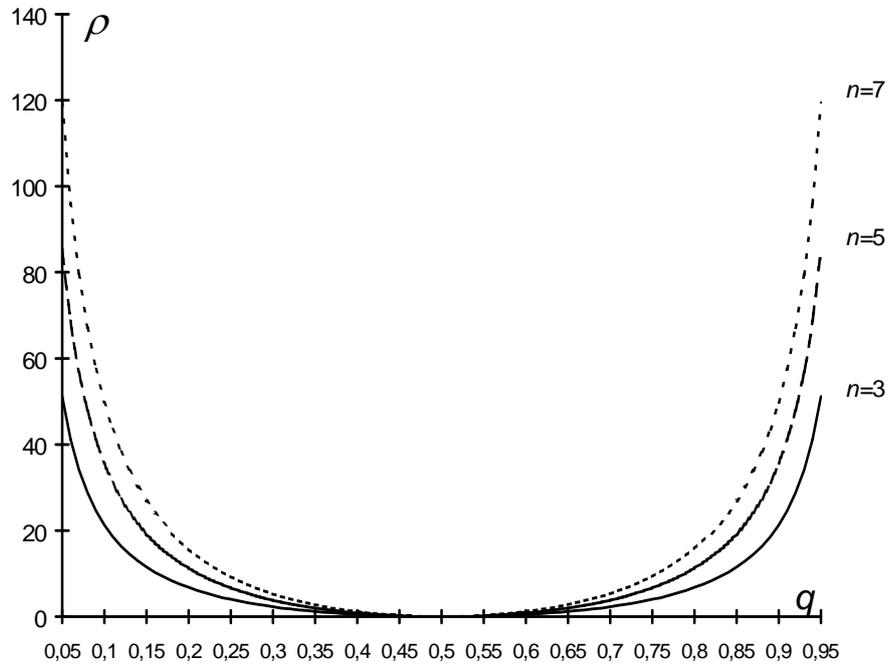


Рис. 9 Семейство зависимостей махаланобисова расстояния от вероятности ошибок на входах мажоритарного элемента для некоторых значений числа входов

Для определения порога $\Theta \equiv a_{n=1}$ рассмотрим рисунок 10, на котором показаны два распределения $F_1(z)$ и $F_2(z)$ случайной величины Z относительно постоянной Θ .

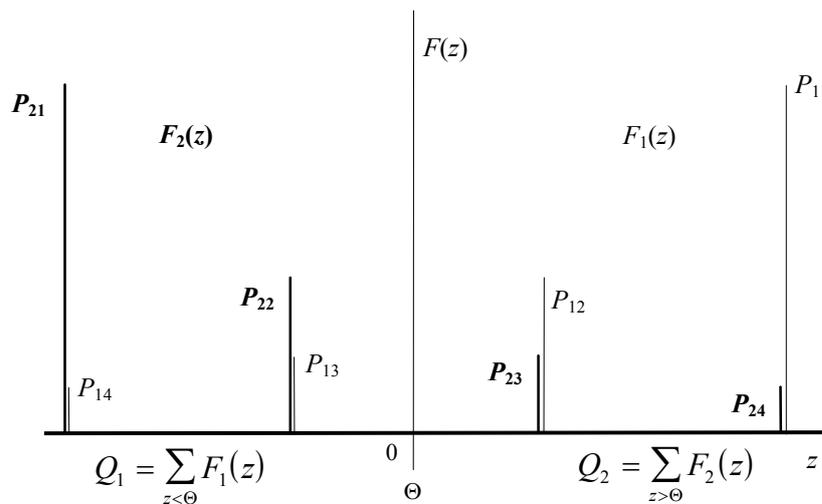


Рис. 10 Дискретные распределения $F_1(z)$ и $F_2(z)$ случайной суммы Z для классов Ω_1 и Ω_2

Если $x = +1$, т.е. если вектор наблюдений \vec{x} принадлежит классу Ω_1 , но $z < \Theta$, то будет принято решение $y = -1$ и произойдет ошибка, вероятность которой

$$Q_1 = \text{Prob}\{Y = -1/x = +1\} = \text{Prob}\{Z < \Theta/x = +1\}$$

определяется выражением

$$Q_1 = \sum_{z < \Theta} F_1(z). \quad (4.33)$$

Если $x = -1$, т.е. если вектор наблюдений \vec{x} принадлежит классу Ω_2 , но реализация z случайной величины Z больше порога Θ ($z > \Theta$), то будет принято решение $y = +1$ и произойдет ошибка, вероятность которой

$$Q_2 = \text{Prob}\{Y = +1/x = -1\} = \text{Prob}\{Z > \Theta/x = -1\}$$

определяется выражением

$$Q_2 = \sum_{z > \Theta} F_2(z). \quad (4.34)$$

Естественно искать такую постоянную Θ , чтобы сумма $Q_1 + Q_2$ вероятностей Q_1 и Q_2 была минимальной. Этому требованию можно удовлетворить, если порог Θ выбрать равноудаленным от математических ожиданий m_1 и m_2 :

$$\Theta = \frac{m_1 + m_2}{2},$$

что на основании соотношения (4.18) даёт для Θ нулевое значение:

$$\Theta = 0.$$

Следовательно, для соотношений (4.33) и (4.34) - с учётом формулы (4.20) - имеем:

$$Q_1 = \sum_{z < 0} F_1(z),$$

$$Q_2 = \sum_{z > 0} F_2(z) = \sum_{z > 0} F_1(-z) = \sum_{-z < 0} F_1(-z) = \sum_{r < 0} F_1(r) = Q_1,$$

где в последнем выражении произведена замена переменной $(-z)$, т.е. индекса суммирования, на r .

Вероятность Q неправильного восстановления сигнала пороговым решающим органом может быть рассчитана по формуле полной вероятности:

$$Q = q_{n+1} \cdot Q_1 + (1 - q_{n+1}) \cdot Q_2 = Q_1 = Q_2, \quad (4.35)$$

где q_{n+1} , как и ранее, является априорной вероятностью предъявления к распознаванию класса Ω_1 , т.е. априорной вероятностью принятия сигналом X значения $x = +1$.

Следовательно,

$$Q = \sum_{z < 0} F_1(z) = \sum_{z < 0} \left(\sum_{i=1}^{n+1} * f_{1i}(z_i) \right), \quad (4.35^*)$$

где $f_{1i}(z_i)$ функция определяется формулой (4.13), а в сумму вносится только тот член, который отвечает отрицательному значению переменной z . Общее же количество дискретных значений переменной z составляет 2^{n+1} , поскольку

$$z = \tilde{a}_1 + \tilde{a}_2 + \dots + \tilde{a}_n + \tilde{a}_{n+1},$$

где \tilde{a}_i может равняться либо $+a_i$, либо $-a_i$.

Каждому дискретному значению переменной z , согласно формуле (1.2.11), соответствует равно произведению некоторых вероятностей слагаемое (1.2.9), т.е. Q_j ($j = \overline{1, n+1}$):

$$Q_j \equiv F_1(z) = \left. \begin{aligned} & \prod_{i=1}^{n+1} f_i(z_i) \equiv \tilde{q}_1 \cdot \tilde{q}_2 \cdot \dots \cdot \tilde{q}_n \cdot \tilde{q}_{n+1} \\ & j = \overline{1, 2^{n+1}} \end{aligned} \right\},$$

где

$$\tilde{q}_k = \begin{cases} q_k, & \text{если } z_k = -a_k \\ 1 - q_k, & \text{если } z_k = +a_k \end{cases}.$$

Для определения вероятности Q достаточно просуммировать те слагаемые Q_j , которые отвечают отрицательным значениям переменной z , т.е.

$$Q = \sum_{z < 0} Q_j = \sum_{z < 0} \tilde{q}_1 \cdot \tilde{q}_2 \cdot \dots \cdot \tilde{q}_n \cdot \tilde{q}_{n+1}. \quad (4.35^{**})$$

Как видим, и полученный этим путём точный алгоритм (4.35**) полностью совпадает с ранее приведённым фундаментальным результатом (1.2.11).

5. Оптимизация весов: байесовский подход

К решению задачи определения порога Θ и весов a_i ($i = \overline{1, n}$) можно предложить и байесовский подход, который в основе своей также является эвристическим. Применительно к проблеме управления он наиболее полно изложен в работе [50].

С этой целью рассмотрим выражения $q_{n+1} \cdot F_1(z)$ и $(1 - q_{n+1}) \cdot F_2(z)$.

Легко видеть, что $q_{n+1} \cdot F_1(z)$ представляет собой ожидаемую вероятность того, что случайная величина Z примет некоторое значение z при $x = +1$. Аналогично, $(1 - q_{n+1}) \cdot F_2(z)$ задаёт ожидаемую вероятность того, что случайная величина Z окажется равной этому же значению z при $x = -1$.

Принятие решения $y = +1$ при условии

$$\sum_{i=1}^n a_i x_i > \Theta \quad (5.1)$$

оправдано лишь в том случае, если

$$\frac{q_{n+1} \cdot F_1(z)}{(1 - q_{n+1}) \cdot F_2(z)} > 1 \quad (5.2)$$

и, наоборот, решение $y = -1$ при

$$\sum_{i=1}^n a_i x_i < \Theta \quad (5.3)$$

следует выносить, когда удовлетворяется неравенство

$$\frac{q_{n+1} \cdot F_1(z)}{(1 - q_{n+1}) \cdot F_2(z)} < 1. \tag{5.4}$$

Легко видеть, что

$$\frac{q_{n+1} \cdot F_1(z)}{(1 - q_{n+1}) \cdot F_2(z)} = \frac{q_{n+1} \cdot f_1(\bar{x})}{(1 - q_{n+1}) \cdot f_2(\bar{x})}. \tag{5.5}$$

Здесь $q_{n+1} \cdot f_1(\bar{x})$ представляет собой ожидаемую вероятность появления вектора \bar{x} наблюдений на входах порогового элемента, когда $x = +1$, а $(1 - q_{n+1}) \cdot f_2(\bar{x})$ задаёт ожидаемую вероятность появления этого же самого вектора при условии, что истинное значение восстанавливаемого сигнала $x = -1$. Следовательно, с учётом соотношения (5.5), неравенства (5.2) и (5.4) примут следующий вид:

$$\frac{q_{n+1} \cdot f_1(\bar{x})}{(1 - q_{n+1}) \cdot f_2(\bar{x})} > 1, \tag{5.2*}$$

$$\frac{q_{n+1} \cdot f_1(\bar{x})}{(1 - q_{n+1}) \cdot f_2(\bar{x})} < 1. \tag{5.4*}$$

Подставляя сюда выражения (4.1) и (4.2), простыми алгебраическими преобразованиями, включающими и логарифмирование, получим:

$$\sum_{i=1}^n x_i \cdot \ln \frac{1 - q_i}{q_i} > \ln \frac{1 - q_{n+1}}{q_{n+1}}, \tag{5.2**}$$

$$\sum_{i=1}^n x_i \cdot \ln \frac{1 - q_i}{q_i} < \ln \frac{1 - q_{n+1}}{q_{n+1}}. \tag{5.4**}$$

Следовательно, если в соотношениях (5.1) и (5.3) положить, что

$$a_i = k \cdot \ln \frac{1 - q_i}{q_i} \Bigg\}_{i=1, n}, \tag{5.6}$$

$$\Theta = k \cdot \ln \frac{1 - q_{n+1}}{q_{n+1}}, \tag{5.7}$$

где k произвольная положительная постоянная, то это не приведёт к нарушению условий (5.2**) и (5.4**), отвечающих оптимальным решениям. Иначе говоря, такой выбор будет согласован с критериями (5.2) и (5.4).

На рисунке 11 для сравнения представлены графики зависимостей (4.29) и (5.6).

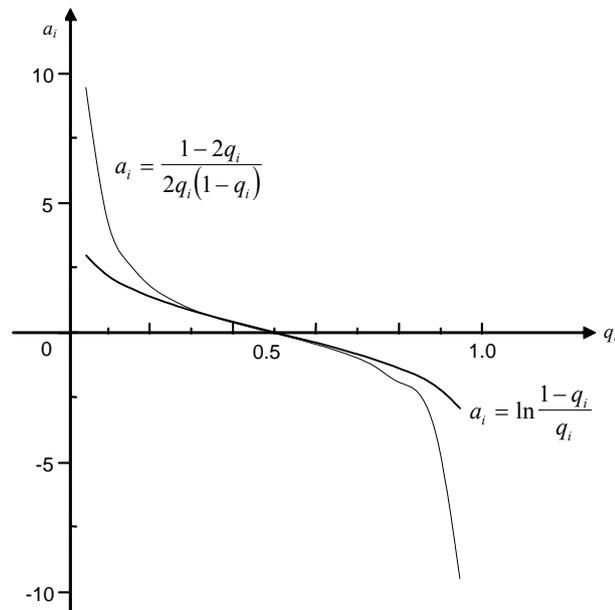


Рис. 11 Зависимость веса a_i от вероятности q_i при $k = 1$, согласно соотношениям (4.29) и (5.6)

Как известно [48], абсолютный предел качества классификации определяется байесовским классификационным правилом. Следовательно, несмотря на одинаковый характер зависимостей (4.29) и (5.6), можно ожидать, что именно стратегии (5.2*) и (5.4*) минимизируют вероятность Q ошибки порогового органа. В принципе это можно было бы показать методами, изложенными в работах [52, 53] Рао и Андерсона. Однако проще исходить из того факта, что вероятность Q ошибки порогового органа определяется вероятностью $Q(e/\bar{x})$ ошибки распознавания заданного образа $\bar{x} = (x_1, x_2, \dots, x_n)'$, принадлежащего классу Ω_i . Здесь, как и прежде, знак «'» означает транспонирование. Эта же вероятность даётся формулой

$$Q(e/\bar{x}) = 1 - P(c/\bar{x}), \quad (5.8)$$

где $P(c/\bar{x})$ - вероятность правильного (корректного) распознавания при тех же условиях.

Для заданного образа \bar{x} , принадлежащего классу Ω_i , имеет место равенство

$$P(c/\bar{x}) = P(\Omega_i/\bar{x}), \quad (5.9)$$

где $P(\Omega_i/\bar{x})$ есть апостериорная вероятность появления класса («гипотезы») Ω_i при условии, что на входах порогового элемента реализовался вектор наблюдений \bar{x} . Следовательно,

$$Q_{\min}(e/\bar{x}) = 1 - P_{\max}(c/\bar{x}) = 1 - \max_i \{P(\Omega_i/\bar{x})\}. \quad (5.10)$$

Но с другой стороны, по формуле Байеса,

$$P(\Omega_1/\bar{x}) = \frac{q_{n+1} \cdot f_1(\bar{x})}{q_{n+1} \cdot f_1(\bar{x}) + (1 - q_{n+1}) \cdot f_2(\bar{x})}, \quad (5.11)$$

$$P(\Omega_2/\bar{x}) = \frac{(1 - q_{n+1}) \cdot f_2(\bar{x})}{q_{n+1} \cdot f_1(\bar{x}) + (1 - q_{n+1}) \cdot f_2(\bar{x})}. \quad (5.12)$$

Следовательно,

$$\frac{q_{n+1} \cdot f_1(\bar{x})}{(1 - q_{n+1}) \cdot f_2(\bar{x})} = \frac{P(\Omega_1/\bar{x})}{P(\Omega_2/\bar{x})}. \quad (5.13)$$

Отсюда следует, что выбор класса Ω_1 по условию (5.2*), а класса Ω_2 по условию (5.4*), автоматически обеспечивает выполнение соотношения (5.10) и минимизирует вероятность $Q_{\min}(e/\bar{x})$ ошибки распознавания заданного образа, принадлежащего классу Ω_i , а, значит, и ожидаемую вероятность Q ошибочной классификации.

Сравнительный анализ рассмотренных двух подходов машинным расчётом вероятности Q ошибки восстановления сигнала показывает, что выбор порога Θ и весов a_i ($i = \overline{1, n}$) по соотношениям (5.7) и (5.6) предпочтительнее.

В частности, при $n = 5, q_6 = 0,5, k = 1, q_1 = 0,1, q_2 = 0,2, q_3 = 0,3, q_4 = 0,4$ и $q_5 = 0,45$ вероятность Q неправильного восстановления сигнала пороговым решающим органом составляет 0,0880, если веса рассчитаны по соотношениям (5.6), и 0,1000, если они выбраны согласно формулам (4.29). Всякое изменение нулевого значения порога в первом случае при неизменных весах ухудшает показатель Q , в то время как во втором случае минимальная вероятность ошибки 0,0951 достигается при значениях Θ , лежащих в интервалах $2 \leq \Theta \leq 2,125$ и $-2,125 \leq \Theta \leq -2$. Однако и при таких значениях порога вероятность неправильного восстановления сигнала остаётся во втором случае выше, чем в первом.

6. Оптимизация весов: энтропийный подход

В решающем органе систему $n + 1$ информационных каналов $B_1, B_2, \dots, B_n, B_{n+1}$ можно трактовать в качестве некоторого источника двоичной информации с энтропией E , определяемой по формуле

$$E = \sum_{i=1}^n E_i, \quad (6.1)$$

где E_i ($i = \overline{1, n+1}$) - энтропия дискретной случайной величины $X \cdot X_i$, распределение которой задаётся в виде совокупности вероятностей q_i и $(1 - q_i)$, отвечающих её реализациям $x \cdot x_i = -1$ и $x \cdot x_i = +1$, соответственно.

Следовательно,

$$E_i = k \cdot [-(1 - q_i) \cdot \ln(1 - q_i) - q_i \cdot \ln q_i], \quad (6.2)$$

где k ($0 < k < \infty$) - произвольная положительная постоянная.

Поэтому

$$E = k \cdot \sum_{i=1}^{n+1} [-(1 - q_i) \cdot \ln(1 - q_i) - q_i \cdot \ln q_i]. \quad (6.3)$$

Если в двух информационных каналах B_i и B_j одинаковые изменения Δq_i и Δq_j ($\Delta q_i = \Delta q_j$) вероятностей их ошибок q_i и q_j влекут за собой различные изменения $(\Delta E)_i$ и $(\Delta E)_j$ энтропии E , то естественно приписать тому информационному каналу,

который вызвал бóльшее изменение энтропии, и бóльший вес. Иначе говоря, вес a_i должен служить мерой изменения энтропии источника информации – совокупности двоичных каналов – в зависимости от приращения вероятности q_i :

$$a_i = \left. \frac{\partial E}{\partial q_i} \right\}_{i=1, n+1}. \quad (6.4)$$

С учётом выражения (6.3) для энтропии E из этой формулы получим:

$$a_i = k \cdot \ln \frac{1-q_i}{q_i}. \quad (6.5)$$

Энтропия дискретной случайной величины, обычно, задаётся в предположении, что $k=1$. Это означает, что в формуле (2.12) соответствующее значение постоянной S , при котором $k=1$, составляет $\exp(-0,5) = 0,60653057$.

7. Связь весов, определённых на основе минимизации махаланобисова расстояния и энтропийного подхода

Обозначим через a_{im} веса, определяемые соотношениями (4.29) и доставляющие обобщённому (махаланобисову) расстоянию максимум:

$$k \cdot \frac{1-2q_i}{2q_i(1-q_i)} = a_{im}. \quad (7.1)$$

Веса же, вычисляемые на основе байесовского подхода, обозначим через a_{ie} :

$$k \cdot \ln \frac{1-q_i}{q_i} = a_{ie}. \quad (7.2)$$

Легко видеть, что

$$a_{im} = \frac{1}{2} \cdot k \cdot \left(\frac{1-q_i}{q_i} - \frac{q_i}{1-q_i} \right). \quad (7.3)$$

С другой стороны

$$\left. \begin{aligned} \frac{1-q_i}{q_i} &= \exp\left(\ln \frac{1-q_i}{q_i}\right) = \exp\left(\frac{1}{k} \cdot a_{ie}\right) \\ \frac{q_i}{1-q_i} &= \exp\left(-\ln \frac{1-q_i}{q_i}\right) = \exp\left(-\frac{1}{k} \cdot a_{ie}\right) \end{aligned} \right\}. \quad (7.4)$$

Учитывая (7.4) в (7.3), получим:

$$a_{im} = k \cdot sh\left(\frac{1}{k} \cdot a_{ie}\right). \quad (7.5)$$

В частности, если $k=1$, то

$$a_{im} = sh(a_{ie}). \quad (7.6)$$

Таким образом, если нормирующий множитель k выбран равным единице, то веса (7.1), доставляющие махаланобисову расстоянию максимум, связаны с весами (7.2), найденными на основе байесовского подхода, монотонным преобразованием (7.6) по закону гиперболического синуса.

8. Восстановление с минимальным риском

Работу решающего органа, приведённого на рисунке 4, представляется возможным описать в терминах теории распознавания образов и с более общих позиций, нежели это было предпринято выше.

В самом деле, значения двоичной переменной X , равные $+1$ и -1 , допустимо рассматривать в качестве классов Ω_1 и Ω_2 , соответственно. Тогда $P(\Omega_1)$ и $P(\Omega_2)$ будут априорными вероятностями появления классов Ω_1 и Ω_2 . Выборочный образ, возникающий на входах решающего элемента и соответствующий истинному значению X , предъявленному к распознаванию, имеет вид последовательности из n положительных и отрицательных единиц, определяющих вектор наблюдений

$$\vec{x} = (x_1, x_2, \dots, x_n)' = (\tilde{1}, \tilde{1}, \dots, \tilde{1})',$$

где тильда над единицей, как и прежде, означает, что последняя берётся либо со знаком плюс, либо со знаком минус. Число различных образов, естественно, составляет 2^n .

Вероятность принадлежности образа \vec{x} классу Ω_i будем обозначать как $P(\Omega_i / \vec{x})$. Иначе говоря, $P(\Omega_i / \vec{x})$ есть апостериорная вероятность появления класса Ω_i при условии, что на входах решающего элемента реализовался вектор наблюдений \vec{x} .

Для описания работы порогового органа в рассмотрение можно ввести матрицу $\|L\|$ размерности 2×2 с элементами L_{ij} . Номера строк этой матрицы могут соответствовать классам Ω_1 и Ω_2 , а номера столбцов – принимаемым решениям $y = y_1 = +1$ и $y = y_2 = -1$:

$$\|L\| = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix}.$$

Элемент L_{ij} этой матрицы $\|L\|$ представляет те убытки (потери), которые несут (терпят) тогда, когда решающий элемент принимает решение о том, что образ \vec{x} принадлежит классу Ω_j , хотя на самом деле он принадлежит классу Ω_i .

Иногда при принятии правильного решения потери равны нулю и одинаковы при принятии любого неправильного решения. Если эти убытки нормировать к единице, элемент L_{ij} матрицы потерь $\|L\|$ можно будет представить в виде

$$L_{ij} = 1 - \delta_{ij},$$

где $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$.

Поскольку образ \vec{x} может принадлежать любому из классов Ω_1 и Ω_2 , в рассмотрение следует ввести математическое ожидание потерь, связанных с отнесением образа \vec{x} к классу Ω_j :

$$R_j(\bar{x}) = \left. \sum_{i=1}^2 L_{ij} \cdot P(\Omega_i / \bar{x}) \right\}_{j=1,2} \quad (8.1)$$

Эта величина в теории распознавания образов часто называется условным средним риском, или условной средней потерей [48, 49].

Работа решающего органа может быть организована следующим образом. Для каждого вектора наблюдений $\bar{x} = (x_1, x_2, \dots, x_n)'$, сформированного набором двоичных каналов B_1, B_2, \dots, B_n , вычисляются условные средние потери $R_1(\bar{x})$ и $R_2(\bar{x})$. Затем решающий элемент причисляет \bar{x} к классу, которому соответствует наименьший условный средний риск.

Легко видеть, что такая стратегия функционирования решающего органа обеспечивает и минимум математического ожидания полных потерь на множестве решений. Решающий орган, минимизирующий математическое ожидание общих потерь, называется байесовским, а принимаемое на его выходе решение – байесовским решением.

Докажем, что пороговый решающий орган на рисунке 4 может реализовать функцию решения с минимальным риском (т.е. дать байесово решение), и найдём необходимые веса и порог.

По формуле Байеса:

$$P(\Omega_i / \bar{x}) = \frac{P(\Omega_i) \cdot P(\bar{x} / \Omega_i)}{P(\bar{x})}, \quad (8.2)$$

где $P(\Omega_i / \bar{x})$ - вероятность формирования образа \bar{x} при условии предъявления к распознаванию класса Ω_i , являющаяся функцией правдоподобия для этого класса, а $P(\bar{x})$ - безусловная ожидаемая вероятность появления вектора наблюдений \bar{x} . Подставляя (8.2) в соотношение (8.1), получим:

$$R_j(\bar{x}) = \left. \frac{1}{P(\bar{x})} \cdot \sum_{i=1}^2 L_{ij} P(\Omega_i) P(\bar{x} / \Omega_i) \right\}_{j=1,2} \quad (8.3)$$

При байесовской стратегии принятия решений образ \bar{x} зачисляется в класс Ω_1 , если $R_1(\bar{x}) < R_2(\bar{x})$, и в класс Ω_2 – в случае противоположного неравенства. При $R_1(\bar{x}) = R_2(\bar{x})$ решение не принимается ($y = 0$).

Нетрудно показать, что условие $R_1(\bar{x}) < R_2(\bar{x})$ равносильно неравенству:

$$L_{11} P(\bar{x} / \Omega_1) P(\Omega_1) + L_{21} P(\bar{x} / \Omega_2) P(\Omega_2) < L_{12} P(\bar{x} / \Omega_1) P(\Omega_1) + L_{22} P(\bar{x} / \Omega_2) P(\Omega_2),$$

откуда

$$(L_{12} - L_{11}) \cdot P(\bar{x} / \Omega_1) P(\Omega_1) > (L_{21} - L_{22}) \cdot P(\bar{x} / \Omega_2) P(\Omega_2).$$

Поскольку $L_{ii} < L_{ij}$, когда $i \neq j$, то последнее соотношение равносильно следующему:

$$\frac{P(\bar{x} / \Omega_1)}{P(\bar{x} / \Omega_2)} > \frac{P(\Omega_2)}{P(\Omega_1)} \cdot \frac{L_{21} - L_{22}}{L_{12} - L_{11}}.$$

В привычных обозначениях метода обобщённого (махаланобисова) расстояния данное неравенство приобретает следующий вид:

$$\frac{f_1(\bar{x})}{f_2(\bar{x})} > \frac{1 - q_{n+1}}{q_{n+1}} \cdot \frac{L_{21} - L_{22}}{L_{12} - L_{11}}.$$

На основании свойства логарифмической функции отсюда следует, что:

$$\ln\left(\frac{f_1(\bar{x})}{f_2(\bar{x})}\right) > \ln\left(\frac{L_{21} - L_{22}}{L_{12} - L_{11}} \cdot \frac{1 - q_{n+1}}{q_{n+1}}\right).$$

Учитывая здесь формулы (4.1) и (4.2), получим:

$$\sum_{i=1}^n x_i \ln \frac{1 - q_i}{q_i} - \ln\left(\frac{L_{21} - L_{22}}{L_{12} - L_{11}} \cdot \frac{1 - q_{n+1}}{q_{n+1}}\right) > 0. \quad (8.4)$$

Полагая, что

$$\left. \begin{aligned} a_i &= k \cdot \ln \frac{1 - q_i}{q_i} \\ i &= \overline{1, n} \\ (0 < k < \infty) \end{aligned} \right\}, \quad (8.5)$$

$$\Theta \equiv a_{n+1} = k \cdot \ln\left(\frac{L_{21} - L_{22}}{L_{12} - L_{11}} \cdot \frac{1 - q_{n+1}}{q_{n+1}}\right) \left. \vphantom{\Theta} \right\}, \quad (8.6)$$

$$0 < k < \infty$$

и принимая во внимание соотношения

$$\left. \begin{aligned} x_{n+1} &\equiv -1 \\ z &= \sum_{i=1}^{n+1} a_i \cdot x_i \end{aligned} \right\}, \quad (8.7)$$

результат, выражаемый неравенством (8.4), можно сформулировать в таком виде: если $z > 0$, то предъявленный к распознаванию образ зачисляется в класс Ω_1 и, следовательно, принимается решение $y = +1$.

Совершенно аналогично преобразования неравенства противоположного смысла приводят к заключению, что, если $z < 0$, то предъявленный к распознаванию образ зачисляется в класс Ω_2 и, следовательно, принимается решение $y = -1$.

Наконец, из условия $R_1(\bar{x}) = R_2(\bar{x})$ можно показать, что, если $z = 0$, то решение не выносится, и, следовательно, $y = 0$.

Три последних результата описывают алгоритм функционирования порогового элемента. Следовательно, байесово решение совпадает с пороговым.

Этим полностью доказывается, что пороговый решающий орган может реализовать функцию решения с минимальным риском, если веса $a_i (i = \overline{1, n})$ выбраны в соответствии с формулой (8.5), а порог $\Theta \equiv a_{n+1}$ - по соотношению (8.6).

9. Законы распределения весов и их статистические характеристики

Приведённые выше результаты показывают, что оптимальный вес a_i существенно зависит от значения вероятности ошибки q_i . Но последняя лишь оценивается в результате M_i независимых испытаний частотой $\omega_i = n_i/M_i$, где n_i - число реализовавшихся ошибок, характеризующихся вероятностью q_i . Разумеется, в длинной серии испытаний теоретически оправдано полагать приближённо, что $q_i = \omega_i$. Однако такое равенство даже при больших M_i сопряжено с некоторой погрешностью, ввиду того, что ω_i есть одно из возможных значений случайной величины $\hat{q}_i = N_i/M_i$, определяемой другой случайной величиной N_i - числом ошибок i -го информационного канала B_i в серии из M_i независимых испытаний. Возможные значения n_i последней составляют дискретный ряд $0, 1, \dots, M_i$.

В силу отмеченного и выбранные веса \hat{a}_i не окажутся оптимальными. В каждой серии испытаний мы будем получать лишь одну из реализаций α_i случайной величины \hat{a}_i , являющейся статистической оценкой оптимального веса a_i . Эта случайная величина в предположении, что константа $k=1$, связана с \hat{q}_i соотношением

$$\hat{a}_i = \ln \frac{1 - \hat{q}_i}{\hat{q}_i} \quad (9.1)$$

где $0 < \hat{q}_i < 1$ и $-\infty < \hat{a}_i < +\infty$.

Естественно поставить задачу определения плотности вероятности $g_i(\alpha_i)$ случайной величины (9.1) по заданной плотности $\varphi_i(\omega_i)$ случайной величины \hat{q}_i .

Как известно, для непрерывной случайной величины \hat{q}_i , имеющей плотность вероятности $\varphi_i(\omega_i)$, и строго монотонной функции $\alpha_i = u(\omega_i)$, плотность вероятности $g_i(\alpha_i)$ случайной величины $\hat{a}_i = u(\hat{q}_i)$ имеет вид:

$$g_i(\alpha_i) = \varphi_i(u^{-1}(\alpha_i)) \cdot \left| \frac{du^{-1}(\alpha_i)}{d\alpha_i} \right|, \quad (9.2.)$$

где $u^{-1}(\alpha_i)$, как обычно, обозначает функцию, обратную функции $\alpha_i = u(\omega_i)$.

Так как в нашем случае

$$\alpha_i = u(\omega_i) = \ln \frac{1 - \omega_i}{\omega_i},$$

поэтому

$$u^{-1}(\alpha_i) = \frac{1}{1 + e^{\alpha_i}},$$

$$\frac{du^{-1}(\alpha_i)}{d\alpha_i} = -\frac{e^{\alpha_i}}{(1 + e^{\alpha_i})^2}.$$

Следовательно,

$$g_i(\alpha_i) = \frac{e^{\alpha_i}}{(1 + e^{\alpha_i})^2} \cdot \varphi_i\left(\frac{1}{1 + e^{\alpha_i}}\right). \quad (9.3)$$

Для задания плотности $\varphi_i(\omega_i)$ случайной величины \hat{q}_i можно использовать байесовский подход. В его основе лежит предположение, что случайная величина N_i имеет плотность распределения $f_i(n_i)$, которая зависит от вероятности q_i ошибки информационного канала B_i .

В традиционном методе статистических выводов q_i является постоянной величиной, а при байесовском подходе q_i трактуется в качестве случайной величины \hat{q}_i и наше априорное знание значений \hat{q}_i описывается плотностью $h_i(\omega_i)$. Эта плотность носит субъективный характер и её нельзя смешивать с объективными оценками вероятности, получаемыми при частотном подходе к расчёту вероятностей.

Проведём M_i независимых опытов, состоящих в том, что i -му информационному каналу B_i для распознавания M_i раз предъявляется двоичная переменная X . В этих испытаниях может быть определена конкретная реализация n_i случайной величины N_i числа ошибок информационного канала. Естественно допустить, что существует условная плотность распределения $G_i(n_i/\omega_i)$ этой случайной величины при данном значении ω_i случайной величины \hat{q}_i . Тогда плотность совместного распределения случайных величин \hat{q}_i и N_i будет иметь вид:

$$f_i(\omega_i, n_i) = h_i(\omega_i) \cdot G_i(n_i/\omega_i). \quad (9.4)$$

Плотность безусловного распределения случайной величины N_i определится по формуле

$$F_i(n_i) = \int_0^1 h_i(\omega_i) G_i(n_i/\omega_i) d\omega. \quad (9.5)$$

Следовательно, плотность условного распределения случайной величины \hat{q}_i при наличии информации о случайной величине N_i примет вид:

$$\varphi_i(\omega_i/n_i) = \left. \begin{aligned} & \frac{h_i(\omega_i) \cdot G_i(n_i/\omega_i)}{F_i(n_i)} \\ & F_i(n_i) > 0 \end{aligned} \right\}. \quad (9.6)$$

Исходя из байесовского подхода, в рассмотрение можно ввести плотность $g_i(\alpha_i/n_i)$ апостериорного распределения случайной величины \hat{a}_i при наличии достоверных данных n_i о значениях N_i . Тогда формула (9.3) примет следующий вид:

$$g_i(\alpha_i/n_i) = \frac{e^{\alpha_i}}{(1+e^{\alpha_i})^2} \cdot \varphi_i\left(\frac{1}{1+e^{\alpha_i}}/n_i\right). \quad (9.7)$$

Поскольку испытания каждого информационного канала проводятся по схеме Бернулли, то случайная величина N_i имеет биномиальное распределение:

$$G_i(n_i/\omega_i) = \left. \begin{aligned} & C_{M_i}^{n_i} \cdot \omega_i^{n_i} \cdot (1-\omega_i)^{M_i-n_i} \\ & n_i = 0, 1, 2, \dots, M_i \end{aligned} \right\}. \quad (9.8)$$

В качестве априорного распределения случайной величины \hat{q}_i примем бета-распределение с плотностью

$$h_i(\omega_i) = \begin{cases} \frac{\omega_i^{a'_i-1} \cdot (1-\omega_i)^{b_i-1}}{B(a'_i, b_i)}, & \text{если } 0 < \omega_i < 1 \\ 0 & \text{при других значениях } \omega_i \end{cases} \quad (9.9)$$

Здесь a'_i и b_i - положительные постоянные, определяющие форму распределения, а $B(a'_i, b_i)$ - т.н. полная бета-функция:

$$B(a, b) = \int_0^1 z^{a-1} (1-z)^{b-1} dz = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

причём

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt.$$

Последнее выражение носит название гамма-функции. Она задаётся для $x > 0$ (в случае же комплексных x гамма-функция определяется для положительных значений действительной части: $\text{Re}(x) > 0$). В случае положительного целого аргумента, как известно,

$$\left. \begin{aligned} \Gamma(n) &= (n-1)! \\ \Gamma(2) &= \Gamma(1) = 1 \end{aligned} \right\}$$

Неполная бета-функция определяется соотношением

$$\left. \begin{aligned} B_x(a, b) &= \int_0^x z^{a-1} (1-z)^{b-1} dz \\ 0 &\leq x \leq 1 \end{aligned} \right\},$$

и, следовательно, $B(a, b) = B_1(a, b)$.

В этих условиях, согласно (9.4),

$$f_i(\omega_i, n_i) = C_{M_i}^{n_i} \cdot \frac{\Gamma(a'_i + b_i)}{\Gamma(a'_i) \cdot \Gamma(b_i)} \cdot \omega_i^{a'_i + n_i - 1} \cdot (1 - \omega_i)^{b_i + M_i - n_i - 1},$$

где $0 < \omega_i < 1$ и $n_i = 0, 1, 2, \dots, M_i$.

Плотность безусловного распределения случайной величины N_i находится по формуле (9.5):

$$F_i(n_i) = C_{M_i}^{n_i} \cdot \frac{\Gamma(a'_i + b_i)}{\Gamma(a'_i) \cdot \Gamma(b_i)} \cdot \int_0^1 \omega_i^{a'_i + n_i - 1} \cdot (1 - \omega_i)^{b_i + M_i - n_i - 1} d\omega_i.$$

Входящий в это выражение сомножителем интеграл представляет собой бета-функцию с параметрами $a'_i + n_i$ и $b_i + M_i - n_i$. Поэтому

$$F_i(n_i) = C_{M_i}^{n_i} \cdot \frac{\Gamma(a'_i + b_i)}{\Gamma(a'_i) \cdot \Gamma(b_i)} \cdot \left. \frac{\Gamma(a'_i + n_i) \Gamma(b_i + M_i - n_i)}{\Gamma(a'_i + b_i + M_i)} \right\}_{n_i = 0, 1, 2, \dots, M_i} \quad (9.10)$$

Данное распределение носит в литературе название гипербиномиального.

Плотность апостериорного распределения случайной величины \hat{q}_i определяется по формуле (9.6) с учётом выражений (9.8), (9.9) и (9.10):

$$\varphi_i(\omega_i/n_i) = C_{M_i}^{n_i} \cdot \frac{\Gamma(a'_i + b_i + M_i)}{\Gamma(a'_i + n_i) \cdot \Gamma(b_i + M_i - n_i)} \cdot \omega_i^{a'_i + n_i - 1} \cdot (1 - \omega_i)^{b_i + M_i - n_i - 1} \cdot \left. \vphantom{\frac{\Gamma(a'_i + b_i + M_i)}{\Gamma(a'_i + n_i) \cdot \Gamma(b_i + M_i - n_i)}}} \right\} \quad (9.11)$$

$$0 < \omega_i < 1$$

Полученное выражение представляет собой бета-распределение с параметрами $a'_i + n_i$ и $b_i + M_i - n_i$. Следовательно, если случайная величина \hat{q}_i имеет априорное бета-распределение, то и апостериорное распределение остаётся бета-распределением, однако параметры a'_i и b_i меняются на $a'_i + n_i$ и $b_i + (M_i - n_i)$ соответственно.

За байесовскую точечную оценку для \hat{q}_i можно взять математическое ожидание случайной величины \hat{q}_i , определив его по соотношению:

$$M[\hat{q}_i] \equiv \mu_{0i} = \int_0^1 \omega_i \varphi_i(\omega_i/n_i) d\omega_i = \frac{a'_i + n_i}{a'_i + b_i + M_i} \quad (9.12)$$

Дисперсия же этой случайной величины вычисляется по формуле

$$D[\hat{q}_i] \equiv \delta_i = \int_0^1 (\omega_i - \mu_{0i})^2 \varphi_i(\omega_i/n_i) d\omega_i = \frac{(a'_i + n_i) \cdot (b_i + M_i - n_i)}{(a'_i + b_i + M_i)^2 \cdot (a'_i + b_i + M_i + 1)} \quad (9.13)$$

Точечная байесовская оценка случайной величины \hat{a}_i может быть определена как её математическое ожидание:

$$M[\hat{a}_i] \equiv \bar{a}_i = \int_0^1 u(\omega_i) \cdot \varphi_i(\omega_i/n_i) d\omega_i, \quad (9.14)$$

где

$$\alpha_i = u(\omega_i) = \ln \frac{1 - \omega_i}{\omega_i}.$$

Точное вычисление \bar{a}_i по формуле (9.14) практически затруднено, так как закон распределения $\varphi_i(\omega_i/n_i)$ случайной величины \hat{q}_i весьма сложен. В этом случае удобно прибегнуть к аппроксимации математического ожидания $M[\hat{a}_i]$ функции $\hat{a}_i = u(\hat{q}_i)$ случайной величины \hat{q}_i .

Для приближённого вычисления математического ожидания \bar{a}_i величины \hat{a}_i последнюю трактуют как функцию случайной величины \hat{q}_i :

$$\hat{a}_i = \ln \frac{1 - \hat{q}_i}{\hat{q}_i} = u(\hat{q}_i).$$

Разложение функции $\hat{a}_i = u(\hat{q}_i)$ в окрестности точки $\hat{q}_i = \mu_{0i}$ в ряд Тейлора до первых трёх членов даёт формулой

$$\hat{a}_i = u(\hat{q}_i) = u(\mu_{0i}) + (\hat{q}_i - \mu_{0i}) \cdot u'(\mu_{0i}) + \frac{(\hat{q}_i - \mu_{0i})^2}{2!} \cdot u''(\mu_{0i}) + R_{0i}, \quad (9.15)$$

где R_{0i} - остаточный член. Математическое ожидание этого выражения имеет вид:

$$\begin{aligned}
 M[\hat{a}_i] &= M[u(\mu_{0i})] + M[\hat{q}_i \cdot u'(\mu_{0i}) - \mu_{0i} \cdot u'(\mu_{0i})] + \\
 &+ M\left[\frac{1}{2} \cdot u''(\mu_{0i}) \cdot (\hat{q}_i - \mu_{0i})^2\right] + M[R_{0i}] = \\
 &= u(\mu_{0i}) + \mu_{0i} \cdot u'(\mu_{0i}) - \mu_{0i} \cdot u'(\mu_{0i}) + \frac{1}{2} \cdot u''(\mu_{0i}) \cdot D[\hat{q}_i] + M[R_{0i}] \approx \\
 &\approx u(\mu_{0i}) + \frac{1}{2} \cdot u''(\mu_{0i}) \cdot D[\hat{q}_i]
 \end{aligned} \quad (9.16)$$

Эта формула является приближённой, так как остаточный член разложения в ряд Тейлора отброшен. Если дисперсия случайной величины \hat{q}_i мала, то можно пренебречь и вторым членом в формуле (9.16). Тогда получим:

$$M[\hat{a}_i] \approx u(\mu_{0i}). \quad (9.17)$$

Следовательно, с учётом формулы (9.1), задающей вид зависимости $\hat{a}_i = u(\hat{q}_i)$, будем иметь:

$$\bar{a}_i = \ln \frac{1 - \mu_{0i}}{\mu_{0i}} + \frac{1}{2} \cdot \frac{1 - 2 \cdot \mu_{0i}}{[\mu_{0i} \cdot (1 - \mu_{0i})]^2} \cdot \delta_i. \quad (9.16')$$

или, в более грубом приближении,

$$\bar{a}_i = \ln \frac{1 - \mu_{0i}}{\mu_{0i}}. \quad (9.17')$$

Учитывая в этих формулах выражения (9.12) и (9.13) для μ_{0i} и δ_i , придём к следующим соотношениям:

$$\bar{a}_i = \ln \frac{b_i + M_i - n_i}{a'_i + n_i} + \frac{1}{2} \cdot \frac{(a'_i + b_i + M_i) \cdot (b_i + M_i - a'_i - 2n_i)}{(a'_i + n_i) \cdot (b_i + M_i - n_i) \cdot (a'_i + b_i + M_i + 1)}. \quad (9.18)$$

В более грубом приближении

$$\bar{a}_i = \ln \frac{b_i + M_i - n_i}{a'_i + n_i}. \quad (9.19)$$

Помимо этого, байесовский подход позволяет рассчитать по формуле (9.7) - с учётом в ней соотношения (9.11) - плотность апостериорного распределения $g_i(\alpha_i/n_i)$, которая выражает степень нашей уверенности относительно значений веса \hat{a}_i с учётом достоверной информации n_i , полученной в серии из M_i независимых опытов по предъявлению i -му информационному каналу B_i для распознавания двоичной переменной X :

$$g_i(\alpha_i/n_i) = \frac{\Gamma(a'_i + b_i + M_i)}{\Gamma(a'_i + n_i) \cdot \Gamma(b_i + M_i - n_i)} \cdot e^{(b_i + M_i - n_i) \cdot \alpha_i} / (1 + e^{\alpha_i})^{(a'_i + b_i + M_i)}. \quad (9.20)$$

Поскольку до опыта известны, строго говоря, лишь границы 0 и 1 области возможных реализаций случайной величины \hat{q}_i , то априорным распределением, доставляющим максимум энтропии этой случайной величины, как известно [103], будет равномерное распределение $h_i(\omega_i)$ с параметрами $a'_i = b_i = 1$.

Для этого частного случая получим:

$$\bar{a}_i = \ln \frac{M_i - n_i + 1}{n_i + 1} + \frac{1}{2} \cdot \frac{(M_i + 2) \cdot (M_i - 2n_i)}{(n_i + 1) \cdot (M_i - n_i + 1) \cdot (M_i + 3)}. \quad (9.21)$$

В более грубом приближении

$$\bar{a}_i = \ln \frac{M_i - n_i + 1}{n_i + 1}. \quad (9.22)$$

Когда при тестировании i -го информационного канала B_i значение n_i оказывается равным нулю, использование формул (9.21) и (9.22) страхует нас от опрометчивого шага назначения бесконечно большого веса \bar{a}_i , что было бы неизбежно при байесовском подходе.

Что касается байесовской плотности распределения случайной величины \hat{a}_i для этого же случая, она принимает следующий вид:

$$g_i(\alpha_i/n_i) = \frac{\Gamma(M_i + 2)}{\Gamma(n_i + 1) \cdot \Gamma(M_i - n_i + 1)} \cdot e^{(M_i - n_i + 1) \cdot \alpha_i} / (1 + e^{\alpha_i})^{(M_i + 2)}. \quad (9.23)$$

Легко показать, что формула (9.19) задаёт моду распределения (9.20), т.е. значение α_i , при котором $g_i(\alpha_i/n_i)$ достигает максимума. Аналогично, выражение (9.22) служит модой распределения (9.23).

10. Краткая сводка результатов и выводы

В статье изучена группа вопросов, связанных с пороговым резервированием на основе модели формального (искусственного) нейрона. Предложен и программно реализован алгоритм расчёта вероятности ошибки порогового органа, исследована проблема назначения входных весов порогового решающего органа, найдены законы распределения этих весов при статистическом разбросе вероятностей ошибок на входах, получено выражение для вероятности неправильного восстановления сигнала мажоритарным элементом для случая, когда число его входов стремится к бесконечности, и, наконец, доказана возможность реализации пороговым элементом функции решения с минимальным риском при соответствующем выборе весов и порога.

Основные выводы, полученные в ней, сводятся к следующему.

1. Точный расчёт вероятности Q неправильного восстановления сигнала пороговым органом с заданными значениями числа входов n , вероятностей их ошибок q_i ($i = \overline{1, n}$), весов a_i ($i = \overline{1, n}$), порога $\Theta \equiv a_{n+1}$ и априорной вероятности q_{n+1} предъявления к распознаванию сигнала +1 осуществляется по формуле

$$Q = \sum_{v < 0} Q_j = \sum_{v < 0} \tilde{q}_1 \cdot \tilde{q}_2 \cdot \dots \cdot \tilde{q}_n \cdot \tilde{q}_{n+1},$$

где

$$v = v_1 + v_2 + \dots + v_n + v_{n+1} = \tilde{a}_1 + \tilde{a}_2 + \dots + \tilde{a}_n + \tilde{a}_{n+1},$$

причём выражение \tilde{a}_i равно либо $+a_i$, либо $-a_i$ и полное число значений v составляет, таким образом, 2^{n+1} . Кроме того, в произведении

$$Q_j = \left. \tilde{q}_1 \cdot \tilde{q}_2 \cdot \dots \cdot \tilde{q}_n \cdot \tilde{q}_{n+1} \right\}_{j=1, 2^{n+1}}$$

суммножитель

$$\tilde{q}_k = \begin{cases} q_k, & \text{если } v_k = -a_k \\ 1 - q_k, & \text{если } v_k = +a_k \end{cases}.$$

Алгоритм имеет удобную для пользователя программную реализацию.

2. Верхняя оценка Q^+ для вероятности ошибки Q порогового элемента имеет вид

$$Q^+ = \prod_{i=1}^{n+1} [(1 - q_i)S^{a_i} + q_i S^{-a_i}],$$

где S – положительное вещественное число, удовлетворяющее условию $0 < S < 1$.

Эта оценка в некоторой степени тривиальна, поскольку при определённых значениях $w_i = S^{a_i}$ отдельные множители

$$Q_i^+ = (1 - q_i)w_i + \frac{q_i}{w_i}$$

превосходят единицу.

3. Минимальная верхняя оценка Q_{\min}^+ выражается формулой

$$Q_{\min}^+ = 2^{n+1} \cdot \prod_{i=1}^{n+1} [\sqrt{q_i(1 - q_i)}],$$

что имеет и следующее эквивалентное представление:

$$Q_{\min}^+ = \exp\left(-\sum_{i=1}^{n+1} A(q_i)\right),$$

где

$$A(q_i) = \left| \ln\left[2\sqrt{q_i(1 - q_i)}\right] \right|.$$

Увеличение числа входов n порогового элемента, если только вероятности ошибок на этих входах не равны $1/2$, ведёт к монотонному уменьшению по экспоненте даже минимальной верхней оценки вероятности ошибочного решения.

4. Значения $w_{oi} = S_0^{a_i}$, доставляющие оценке Q^+ минимум Q_{\min}^+ , удовлетворяют условию

$$w_{oi} = S_0^{a_i} = \sqrt{\frac{q_i}{1 - q_i}}, \quad (i = \overline{1, n+1}).$$

Если веса a_i выбраны в соответствии с этим условием по формуле

$$a_i = \left. \frac{1}{2|\ln S_0|} \cdot \ln \frac{1 - q_i}{q_i} \right\}_{i = \overline{1, n+1}},$$

где $0 < S_0 < 1$, то при заданных n и $q_i (i = \overline{1, n+1})$ вероятность Q ошибки порогового элемента не будет превосходить величины Q_{\min}^+ .

5. Из эвристических соображений вытекает возможность выбора таких весов $a_i (i = \overline{1, n})$, при которых математические ожидания m_1 и m_2 случайной суммы $Z = \sum_{i=1}^n a_i X_i$ при условии предъявления к распознаванию сигналов $X = +1$, и $X = -1$, соответственно, максимально удалены друг от друга, а дисперсия σ_z^2 этой суммы минимальна. Целевой функцией при таком подходе служит обобщённое расстояние Махаланобиса

$$\rho = \frac{(m_1 - m_2)^2}{\sigma_z^2}.$$

Веса $a_i (i = \overline{1, n})$, доставляющие этому выражению максимум, удовлетворяют условию

$$a_i = k \left. \begin{array}{l} \frac{1 - 2q_i}{2q_i(1 - q_i)} \\ i = \overline{1, n} \end{array} \right\},$$

где $k (0 < k < \infty)$ – произвольная положительная константа.

6. Веса $a_i (i = \overline{1, n+1})$, рассчитываемые на основе байесовского подхода (по критерию максимальной апостериорной вероятности), задаются формулами:

$$a_i = k \cdot \left. \begin{array}{l} \ln \frac{1 - q_i}{q_i} \\ i = \overline{1, n+1} \\ (0 < k < \infty) \end{array} \right\}.$$

7. Если совокупность $n + 1$ информационных каналов $B_1, B_2, \dots, B_n, B_{n+1}$ рассматривать как некоторый единый источник двоичной информации

$$\vec{x} = (x_1, x_2, \dots, x_n, x_{n+1})' = (\tilde{1}, \tilde{1}, \dots, \tilde{1}, -1)',$$

то энтропия этого источника

$$E = k \cdot \sum_{i=1}^{n+1} [-(1 - q_i) \cdot \ln(1 - q_i) - q_i \cdot \ln q_i].$$

Вес a_i информационного канала $B_i (i = \overline{1, n+1})$ следует определить как меру изменения энтропии E в зависимости от вариации вероятности q_i его ошибки:

$$a_i = \left. \frac{\partial E}{\partial q_i} \right\}_{i = \overline{1, n+1}}.$$

На основании этого определения при энтропийном подходе к проблеме назначения оптимальных весов имеем:

$$a_i = k \cdot \ln \frac{1-q_i}{q_i} \Bigg\}_{i=1, n+1}.$$

8. Сравнивая этот результат с предыдущим, можно утверждать, что веса a_i , определяемые в соответствии с формулой $a_i = \frac{\partial E}{\partial q_i}$, являются согласованными и с критерием максимальной апостериорной вероятности. Это означает, что, если веса a_i ($i = \overline{1, n+1}$) выбраны согласно энтропийному подходу, решающий элемент порогового типа будет выдавать такое двоичное решение, вероятность безошибочности которого больше вероятности безошибочности противоположного решения.

9. Если нормирующие множители весов выбрать равными единице, то веса

$$a_{im} = \frac{1-2q_i}{2q_i(1-q_i)},$$

доставляющие обобщённому (махаланобисову) расстоянию максимум, связаны с весами

$$a_{ie} = \ln \frac{1-q_i}{q_i},$$

согласованными с критерием максимальной апостериорной вероятности, монотонным преобразованием по закону гиперболического синуса:

$$a_{im} = sh(a_{ie}).$$

10. Для мажоритарного решающего элемента (порогового элемента со следующими данными: $a_{n+1} \equiv \Theta = 0$, $a_i = 1$, $q_i = q$ для всех $i = \overline{1, n}$) минимальная верхняя оценка вероятности ошибки на его выходе имеет следующий вид:

$$Q_{\min}^+ = \left[2 \cdot \sqrt{q(1-q)} \right]^n,$$

или, что то же самое,

$$Q_{\min}^+ = \exp[-n \cdot A(q)],$$

где

$$A(q) = \left| \ln 2 \sqrt{q(1-q)} \right|.$$

Эта оценка справедлива при условии, что $q_{n+1} = 1/2$ и $0 < q < 1$, так как только в этом случае постоянная

$$S_0 = \sqrt{q/(1-q)}$$

удовлетворяет требованию $0 < S_0 < 1$, а порог $\Theta \equiv a_{n+1}$.

11. Аналитическая оценка для верхней границы Q_a^+ вероятности ошибки на выходе мажоритарного элемента при биномиальном распределении числа ошибок на n его входах, когда $n \rightarrow \infty$, совпадает с приведённой в предыдущем пункте минимальной верхней оценкой, т.е.

$$Q_a^+ = \exp[-n \cdot A(q)],$$

если только

$$0 < q < q_0(n).$$

Здесь

$$q_0(n) = \frac{\left\lfloor \frac{n}{2} + 1 \right\rfloor + 1}{n + 1},$$

где $\left\lfloor \frac{n}{2} + 1 \right\rfloor$ – наибольшая целая часть величины $(n/2 + 1)$, причём

$$\lim_{n \rightarrow \infty} q_0(n) = \frac{1}{2}.$$

12. Этот результат, равно как и приведённый в пункте 3 настоящей сводки, обнаруживает глубокую внутреннюю связь с теоремой Шеннона, которая утверждает, что число сообщения заданной длины n (длительности τ), составленных из отдельных символов – как в отсутствие так и при наличии фиксированных и вероятностных ограничений (в последнем случае при условии эргодичности источника) – с увеличением n (или τ) возрастает по асимптотически показательному закону. Упомянутая связь может быть сформулирована в виде следующего положения: по какому закону с увеличением n двоичных каналов возрастает на входах решающего элемента количество той информации, на основе которой может приниматься решение y , по такому же (степенному) закону убывает при этом вероятность того, что принятое решение ошибочно.

13. Оптимальные веса a_i зависят от значений вероятности ошибки q_i (при фиксированном нормирующем множителе k). Однако в экспериментах определяются лишь возможные значения случайной величины \hat{q}_i , являющейся оценкой вероятности q_i . В силу этого в каждой серии испытаний удаётся получить лишь одну из реализаций α_i случайной величины $\hat{\alpha}_i$, являющейся статистической оценкой веса α_i и связанной с \hat{q}_i соотношением

$$\hat{a}_i = \ln \frac{1 - \hat{q}_i}{\hat{q}_i},$$

если предположить, что $k = 1$. Здесь $0 < \hat{q}_i < 1$ и $-\infty < \hat{a}_i < +\infty$. Решение задачи определения плотности вероятности $g_i(\alpha_i)$ случайной величины \hat{a}_i по заданной плотности $\varphi_i(\omega_i)$ случайной величины \hat{q}_i даётся формулой

$$g_i(\alpha_i) = \frac{e^{\alpha_i}}{(1 + e^{\alpha_i})^2} \cdot \varphi_i\left(\frac{1}{1 + e^{\alpha_i}}\right).$$

14. Исходя из байесовского подхода, предпочтительнее ввести в рассмотрение плотность $g_i(\alpha_i/n_i)$ апостериорного распределения случайной величины \hat{a}_i при наличии достоверных данных n_i о значении случайного числа N_i ошибок i -го информационного канала B_i в серии из M_i независимых испытаний. В частности, если случайная величина \hat{q}_i имеет априорное бета-распределение с параметрами a'_i и b_i , указанная плотность определяется формулой:

$$g_i(\alpha_i/n_i) = \frac{\Gamma(a'_i + b_i + M_i)}{\Gamma(a'_i + n_i) \cdot \Gamma(b_i + M_i - n_i)} \cdot e^{(b_i + M_i - n_i) \cdot \alpha_i} / (1 + e^{\alpha_i})^{(a'_i + b_i + M_i)}.$$

15 Математическое ожидание \bar{a}_i случайной величины \hat{a}_i имеет для этого случая следующий вид:

$$\bar{a}_i = \ln \frac{M_i - n_i + 1}{n_i + 1} + \frac{1}{2} \cdot \frac{(M_i + 2) \cdot (M_i - 2n_i)}{(n_i + 1) \cdot (M_i - n_i + 1) \cdot (M_i + 3)}.$$

16. Поскольку до опыта известны, строго говоря, лишь границы 0 и 1 области возможных реализаций величины \hat{q}_i , то априорным распределением, доставляющим максимум энтропии этой случайной величины, будет равномерное распределение, т.е. бета-распределение с параметрами $a'_i = b_i = 1$. Для этого частного случая

$$\bar{a}_i = \ln \frac{M_i - n_i + 1}{n_i + 1} + \frac{1}{2} \cdot \frac{(M_i + 2) \cdot (M_i - 2n_i)}{(n_i + 1) \cdot (M_i - n_i + 1) \cdot (M_i + 3)}.$$

17. Пороговый решающий орган может реализовать байесовскую стратегию принятия решения, если веса $a_i (i = \overline{1, n})$ его n входов выбраны по соотношению

$$\left. \begin{array}{l} a_i = k \cdot \ln \frac{1 - q_i}{q_i} \\ i = \overline{1, n} \\ (0 < k < \infty) \end{array} \right\},$$

а порог $\Theta \equiv a_{n+1}$ – в соответствии с формулой

$$\Theta = k \cdot \ln \left(\frac{L_{21} - L_{22}}{L_{12} - L_{11}} \cdot \frac{1 - q_{n+1}}{q_{n+1}} \right),$$

где $q_i (i = \overline{1, n})$ – вероятность ошибки i -го входа; q_{n+1} – априорная вероятность предъявления к распознаванию класса Ω_1 , т.е. двоичного сигнала $X = +1$; $(1 - q_{n+1})$ – априорная вероятность предъявления к распознаванию класса Ω_2 , т.е. двоичного сигнала $X = -1$; L_{ij} – элемент матрицы потерь $\|L\|$.

В этой матрице L_{ij} представляет те убытки, которые несут, принимая решение о принадлежности образа \vec{x} классу Ω_j , когда на самом деле он принадлежит классу Ω_i .

При байесовской стратегии принятия решений выборочный образ $\vec{x} = (x_1, x_2, \dots, x_n)'$, появившийся на входах порогового элемента, зачисляется в класс Ω_j по критерию минимума среднего риска $R_j(\vec{x})$, задающего математическое ожидание потерь, связанных с отношением вектора наблюдений \vec{x} к классу Ω_j .

Литература

1. Нейман Дж. Вероятностная логика и синтез надёжных организмов из ненадёжных компонент // Автоматы: Пер. с англ.-М.: ИЛ, 1956.-С.68-139.
2. Бебиашвили Ш.Л. Основные вопросы теории резервирования // Известия АН СССР: ОТН.-1956.-№2.-С.69-74.
3. Мур Э.Ф., Шеннон К.Э. Надёжные схемы их ненадёжных реле // Кибернетический сборник: Пер. с англ.-М.: ИЛ, 1960.-Вып.1.-С.109-148.
4. Miller R. Majority Logic Analysis: Publ. ¹ M-895 / Hermes Electr. Co.-Cambridge (Mass.), 1960.
5. Wasserman R., Brown W.G., Tierney J. Improvements of Electronic Computer Reliability through the use of Redundancy // Proc. Nat. Electron. Conf.-Chicago, 1961/-V.17.-P.341-359.
6. Tryon J.G. Redundant Logic Circuitry: U. S. Patent 2, 942,193.-1958.
7. Трайон Дж. Учетверённая логика // Методы введения избыточности для вычислительных машин: Пер. с англ.-М.: Сов. радио, 1966.-С.241-265.
8. Pierce W.H. A Proposed System of Redundancy to Improve the Reliability of Digital Computers: Tech. Rep. `1552-1/ Stanford Univ., Solid State Elektron. Lab.-1960.
9. Pierce W.H. Improving Reliability of Digital Systems by Redundancy and Adaptation: Tech. Rep.`1552-3/Stanford Univ., Solid State Electron.Lab.-1962.
10. Пирс У. Адаптивные органы голосования повышают эффективность избыточности // Методы введения избыточности для вычислительных машин: Пер. с англ.-М.: Сов. радио, 1966.-С.266-288.
11. Pierce W.H. Adaptive Decision Elements to Improve the Reliability of Redundant Systems // IRE Intern. Convention Record.-1962.-V.10.-Pt 4.-P.124-131.
12. Пирс У. Построение надёжных вычислительных машин: Пер. с англ.-М.: Мир, 1968.-270 с.
13. Лоуэншусс О. Восстанавливающие органы в избыточных автоматах // Кибернетический сборник: Пер. с англ.-М.: ИЛ, 1961.-Вып.2.-С.206-228.
14. Mann W.C. Systematically Introduced Redundancy in Logical Systems // IRE Intern. Convention Record.-1961.-V.9.-Pt 2.-P.241-263.
15. Манн У. Процессы восстановления в избыточных вычислительных системах // Методы введения избыточности для вычислительных машин: Пер. с англ.-М.: Сов. радио, 1966.-С.309-330.
16. Левин В.И. Элементы матричной теории автоматов.-Рига: Зинатне, 1973.-231 с.
17. Armstrong D.B. A General Method of Applying Error Correctio to Synchronous Digital Systems// Bell System Tech. J.-1961.-V.40.-P.577-597.
18. Хетагуров Я.А., Руднев Ю.П. Повышение надёжности цифровых устройств методами избыточного кодирования.-М.: Энергия, 1974.-271 с.
19. Сагалович Ю.Л. Кодирование состояний и надёжности автоматов.-М.: Связь, 1975.-205 с.
20. Хэмминг Р.В. Коды с обнаружением и исправлением ошибок // Коды с обнаружением и исправлением ошибок: Пер. с англ.-М.: ИЛ, 1956.-С.7-22.
21. Возенкрафт Дж. Рейффен Б. Последовательное декодирование: Пер. с англ.-М.: ИЛ, 1963.-153 с.
22. Диллон Б., Сингх Ч. Инженерные методы обеспечения надёжности систем: Пер. с англ.-М.: Мир, 1984.-318 с.
23. Вербик Л.М. Минимизация ошибки нейронных сетей // Принципы самоорганизации: Пер. с англ.-М.: Мир, 1966.-С.163-177.
24. Варшавский В.И. Некоторые вопросы теории логических сетей, построенных из пороговых элементов // Вопросы теории математических машин.-М.: Физматгиз, 1962.-Вып.2.-С.52-106.
25. Варшавский В.И. Функциональные возможности и синтез пороговых элементов // Докл. АН СССР.-1961.-Т.139.-№5.-С.1071-1074.
26. Mc Naughton R. Unate Truth Functions//IRE Trans. on Electr. Computers.-1961.-V.10.-P.1-6.
27. Paull M.C., Mc Cluskey E.J. Boolean Functions Realizable with Single Threshold Devices // Proc. IRE.-1960.-V.48.-P.1335-1337.
28. Stram O.B. Arbitrary Boolean Functions of N Variable in Terms of Threshold Devices // Proc. IRE.-1961.-V.49.-P.210-220.
29. Дертоузос М. Пороговая логика: Пер. с англ.-М.: Мир, 1966.-344 с.
30. Боголюбов И.Н., Овсевич Б.Л., Розенблюм Л.Я. Синтез схем из пороговых и мажоритарных элементов // Сети передачи информации и их автоматизация.-М.: Наука, 1965.-С.80-126.
31. Закревский А.Д. Методы синтеза функционально устойчивых автоматов // Докл. АН СССР.-1959.-Т.129.- №4.-С.729-731.

32. Кон М., Линдеман Р. Аксиоматическая теория пороговой логики // Экспресс-информация: Вычислительная техника.-1961.- №30.-Реф.153.-С.1-11.
33. Вавилов Е.Н., Егоров Б.М., Ланцев В.С., Тоценко В.Г. Синтез схем на пороговых элементах.-М.: Сов. радио, 1970.-368 с.
34. Грибанов Г.Н. Построение избыточных однотактных логических устройств с мажоритарным декодированием // Труды Уральского ордена Трудового Красного Знамени политехнического института им. С.М. Кирова.-1968.- №168.-С.119-124.
35. Доманицкий С.М. Построение надёжных логических устройств.-М.: Энергия, 1971.-279 с.
36. Kuehn R.E. Computer Redundancy, Design, Performance and Future // IEEE Trans. Reliab.-1969.-V.R-18.-P.87-96.
37. Пакулов Н.И., Уханов В.Ф., Чернышов П.Н. Мажоритарный принцип построения надёжных узлов и устройств ЦВМ.-М.: Сов. радио, 1974.-184 с.
38. Харкевич А.А. Опознавание образов: Избранные труды.-М.: Наука, 1973.-Т.3.-С.456-468.
39. Себестиан Г.С. Процессы принятия решений при распознавании образов: Пер. с англ.-Киев: Техника, 1965.-151 с.
40. Горелик А.Л., Скрипник В.А. Методы распознавания.-М.: Высшая школа, 1977.-220- с.
41. Васильев В.И. Распознающие системы.-Киев: Наукова думка, 1969.-292 с.
42. Фу К. Последовательные методы в распознавании образов и обучении машин: Пер. с англ.-М.: Наука, 1971.-156 с.
43. Загоруйко Н.Г. Методы распознавания и их применение.-М.: Сов. радио, 1972.-206 с.
44. Барабаш Ю.Л. Коллективные статистические решения при распознавании.-М.: Радио и связь, 1983.-224 с.
45. Растринин Л.А., Эринштейн Р.Х. Метод коллективного распознавания.-М.: Энергоиздат, 1981.-80 с.
46. Ховард Р.А. Анализ решений: применение при статистическом выводе, принятии решения и планировании эксперимента // Труды института инженеров по электротехнике и радиоэлектронике – ТИИЭР: Пер. с англ.-Мир, 1970.-Т.58.-№5.-С.32-45.
47. Фукунага К. Введение в статистическую теорию распознавания образов: Пер. с англ.-М.: Наука, 1979.-369 с.
48. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер. с англ.-М.: Мир, 1978.-412 с.
49. Дуда Р., Харт П. Распознавание образов и анализ сцен: Пер. с англ.-Мир, 1976.-511 с.
50. Моррис У. Наука об управлении. Байесовский подход: Пер. с англ.-Мир, 1971.-304 с.
51. Фишер Р. Статистические методы для исследователей: Пер. с англ.-М.: Госстатиз, 1958.-268 с.
52. Рао С.Р. Линейные статистические методы и их применение: Пер. с англ.-М.: Наука, 1968.-574 с.
53. Андерсон Т. Введение в многомерный статистический анализ: Пер. с англ.-М.: Физматгиз, 1968.-500 с.
54. Аффифи А., Эйзен С. Статистический анализ. Подход с использованием ЭВМ: Пер. с англ.-М.: Мир, 1982.-488 с.
55. Мультипроцессорные системы и параллельные вычисления: Пер. с англ. / под ред. Ф.Г. Энслоу.-М.: Мир, 1976.-383 с.
56. Пospelов Д.А. Введение в теорию вычислительных систем.-М.: Сов. радио, 1972.-280 с.
57. Анисимов Б.В., Четвериков В.Н. Основы теории и проектирования цифровых вычислительных машин.-М.: Высшая школа, 1970.-575 с.
58. Балашов Е.П., Кноль А.И. Многофункциональные запоминающие устройства.-Л.: Энергия, 1972.-143 с.
59. Варшавский В.И. и др. Однородные структуры.-М.: Энергия, 1973.-151 с.
60. Евреинов Э.В. Однородные вычислительные системы, структуры и среды.-М.: Радио и связь, 1981.-208 с.
61. Евреинов Э.В., Прангишвили И.В. Цифровые автоматы с настраиваемой структурой (Однородные среды).-М.: Энергия, 1971.-240 с.
62. Балашов Е.П., Смоллов В.Б., Петров Г.А., Пузанков Д.В. Многофункциональные регулярные вычислительные структуры.-М.: Сов. радио, 1978.-288 с.
63. Прангишвили И.В. и др. Однородные микроэлектронные ассоциативные процессоры.-М.: Сов. радио, 1973.-280 с.
64. Прангишвили И.В. и др. Микроэлектроника и однородные структуры для построения логических и вычислительных устройств.-М.: Наука, 1967.-280 с.
65. Кантария Г.В. Параллельное микропрограммирование и принципы проектирования устройств управления ЦВМ // Изв. АН СССР, ОТН, Энергетика и автоматика.-1962.-№5.-С140-143.

66. Кантария Г.В. Оптимальный двухуровневый компромиссно-согласованный выбор // Сообщения АН Груз. ССР.-1976.-Т.81.- №3.-С.581-584.
67. Кантария Г.В. Оптимальный выбор с согласованием альтернативных гипотез // Сообщения АН Груз. ССР.-1972.-Т.68.- №3.-С.553-556.
68. Цирамуа Г.С. Основы теории адаптивно перестраиваемых вычислительных систем на базе многофункциональных элементов / Автореферат диссертации на соискание учёной степени доктора технических наук.-Ленинград, 1980.-49 с.
69. Левин В.И. Вероятностные методы исследования надёжности и точности конечных автоматов / Автореферат диссертации на соискание учёной степени доктора технических наук.-Рига, 1971.-44 с.
70. Потапов В.И. Анализ и синтез высоконадёжных цифровых вычислительных и логических структур из пороговых модулей.-Новосибирск, 1977.-76 с.
71. Доманицкий С.М. Методы построения избыточных структур с восстанавливающими органами.-Киев, 1968.-35 с.
72. Курочкин О.А., Смирнов А.С. О построении самовосстанавливаемых структур // Автоматика и телемеханика.-1984.- №1.-С.135-138.
73. Курочкин О.А., Смирнов А.С., Степанов В.А. Информационная и функциональная надёжность дискретных устройств // Автоматика и телемеханика.-1983.- №6.-С.163-168.
74. Закревский А.Д. Функциональная устойчивость релейных схем // Труды Сибирского физико-технического института.-Томск: Изд-во Томского университета.-1961.-Вып.40.-170 с.
75. Францис Т.А., Янбых Г.Ф. Избыточность в электронных дискретных устройствах.-Л.: Энергия, 1969.-248 с.
76. Руднев Ю.П. Оценка степени улучшения надёжности при применении корректирующих кодов // Кибернетика.-1966.-№3.-С.67-71.
77. Гаврилов М.А. Структурная избыточность и надёжность работы релейных устройств // Труды 1-го международного конгресса ИФАК.-М.: Изд-во АН СССР, 1961.-Т.3.-108 с.
78. Церцвадзе Г.Н. Стохастические автоматы и задачи построения надёжного автомата из ненадёжных элементов // Автоматика и телемеханика.-1964.- №2.-С.347-353.
79. Скляревич А.Н. Надёжность систем с накоплением нарушений.-Рига: Зинатне, 1969.-211 с.
80. Пархоменко П.П, Согомонян Е.С. Основы технической диагностики: Оптимизация алгоритмов диагностирования. Аппаратурные средства.-М.: Энергоиздат, 1981.-319 с.
81. Гройсберг Л.Б., Линденбаум М.Д. Расчёт надёжности систем с произвольной структурой при общем кворумном резервировании // Техническая кибернетика.-1974.-№2.-С.66-70.
82. Митичкин Е.Н. Учёт функциональной связи элементов при расчёте мажоритарной структуры // Техническая кибернетика.-1975.-№5.-С.124-128.
83. Доманицкий С.М. Расчёт надёжности логических элементов и одноктактных систем управления // Автоматика и телемеханика.-1965.-№5.-С.898-905.
84. Черкесов Г.Н. Надёжность технических систем с временной избыточностью.-М.: Сов. радио, 1974.-295 с.
85. Шишонок Н.А. Эксплуатационная работоспособность систем длительного пользования и потребный ЗИП // Лекции по теории надёжности.-Киев: Изд-во КВИРТУ, 1968.-110 с.
86. Логинов В.И. Реализация логических функций на одном пороговом элементе // Изв. АН СССР, Техническая кибернетика.-1966.- №1.-С.35-41.
87. Гурьев Д.И. Нейронные сети Хопфилда и сложные логические задачи // Приборостроение.-1994.- №3-4.-С.63-73.
88. Lippman R. P. «An Introduction to Computing with Neural Nets» // IEEE ASSP Magazine.-1987.-V.4.- N4.- P.4-22.
89. Hakan Özdemir (Member, IEEE), Asım Kepkep (Student Member, IEEE), Banu Pamir (Student Member, IEEE), Yusuf Leblebici (Member, IEEE) and Uğur Çilingiroğlu (Member, IEEE) - A Capacitive Threshold-Logic Gate, 1996.
90. Gogiashvili J.G., Namicheishvili O.M., Chonia G.G. Optimization of Weights for Threshold Redundancy of Binary Channels by the Method of (Mahalanobis') Generalised Distance. MMR'2000 - Second International Conference on Mathematical Methods in Reliability: Methodology, Practice and Interference; Université Victor Segalen Bordeaux 2; Bordeaux, France, July 4-7, 2000; Abstracts' Book, V.1, P. 463-466.

91. J. M. Quintana, M. J. Avedillo, J. L. Huertas Efficient Realization of a Threshold Voter for Self-Purging Redundancy Journal of Electronic Testing: Theory and Applications Volume 17 , Issue 1 (February 2001) Pages: 69 – 73.
92. Alexander B. Romanovsky (State Technical University, St.Petersburg, Russia) - Fault Tolerance: Synchronization of Redundancy, 2001.
93. Myron Hecht (SoHaR Incorporated), Herbert Hecht (SoHaR Incorporated), Eltefaat Shokri (Compaq Computer) - Adaptive Fault Tolerance for Spacecraft, 2000.
94. Gil Kalai (Hebrew University of Jerusalem and Yale University) and Shmuel Safra (Tel Aviv University, Princeton University, and Institute for Advanced Study, Princeton) - Threshold Phenomena and Influence with Some Perspectives from Mathematics, Computer Science, and Economics, 2005.
95. Steve Ferrera, Nicholas P. Carter (Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, Illinois, USA) - Reconfigurable Circuits Using Hybrid Hall Effect Devices, 2001.
96. Dong-Wan Ryoo, Jeun-Woo Lee - Implementation of Adaptive Filter with Fault Tolerance, 2004.
97. J.M. Quintana, M.J. Avedillo and J.L. Huertas (Instituto de Microelectrónica de Sevilla, IMSE-CNM, Univ. de Sevilla) - Realization of a Threshold Voter for Self-Purging Redundancy, 2001.
98. Yauheni B. Veryha (Department of Machine Building, Belarussian State Polytechnical Academy, Minsk, Republic of Belarus) Neural Network Error Accommodation with Fuzzy Logic Elements in Robot Time Optimal Path Tracking, 2001.
99. Peter Jung and Gottfried Mayer-Kress (Department of Physics and Center for Complex Systems Research University of Illinois at Urbana-Champaign, Urbana) - Stochastic Resonance in Threshold Devices, 2004.
100. Rui Zhang, Pallav Gupta, Lin Zhong and Niraj K. Jha (Department of Electrical Engineering, Princeton University) - Synthesis and Optimization of Threshold Logic Networks with Application to Nanotechnologies, 2005.
101. Арчил Прангишвили, Олег Намичейшвили, Арчил Элизбарашвили Нейронные сети.-Тбилиси: Изд-во Технический университет, 2007.-164 с. (на груз. языке)
არჩილ ფრანგიშვილი, ოლეგ ნამიჩეიშვილი, არჩილ ელიზბარაშვილი ნეირონული ქსელები.-თბილისი: გამომცემლობა «ტექნიკური უნივერსიტეტი», 2007.-164 გვ.
102. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров: Пер. с англ.- М.: Наука.-1977.-832 с.
103. Голдман С. Теория информации: Пер. с англ.-М.: ИЛ.-1957.-446 с.
104. Mahalanobis P.C. On the Generalized Distance in Statistics // Proceedings of the National Institute of Sciences of India.-1936.-V.12.-P.49-55.

Статья получена: 2009-04-30