

Т. Шарфельд
Системы RFID низкой стоимости

с Приложениями И. Девиля, Ж. Дамура, Н. Чаркани, С. Корнеева и А. Гуларии. Перевод с английского и научная редакция С. Корнеева.

Москва 2006

Аннотация. Излагается теория, рассматриваются вопросы проектирования и практического использования аппаратуры, которая реализует технологию RFID. Детально обсуждаются физические, технические и административные ограничения, в рамках которых возможно создание аппаратуры и, в частности, пассивных меток RFID низкой стоимости.

Книга будет полезна студентам радиотехнических специальностей, аспирантам и специалистам – системным интеграторам и разработчикам, желающим совершенствовать свой опыт в области перспективной технологии радиочастотной идентификации.

Содержание

Предисловие	5
1. Введение	6
1.1. Вводные замечания.....	6
1.2. Краткий исторический обзор.....	8
1.3. Классификация систем RFID.....	9
1.3.1. Чиповые и бесчиповые метки.....	10
1.3.2. Пассивные, полуактивные и активные метки.....	11
1.3.3. ROM и WORM метки.....	11
1.4. Предположения.....	12
1.5. Характеристики аппаратуры.....	12
1.5.1. Технические параметры.....	13
1.5.2. Стоимость и размеры.....	14
1.6. Функции, ограничения и структура работы.....	15
1.6.1. Компоненты и их функции.....	15
1.6.2. Ограничения и структура работы.....	17
2. Электродинамика и антенны	19
2.1. Введение.....	19
2.2. Уравнения Максвелла и основы электродинамики.....	19
2.3. Антенны и окружающее пространство.....	21
2.4. Импедансы пространства, антенн и цепей.....	24
2.5. Связь в дальней и ближней зоне поля.....	26
2.5.1. Связь в ближней зоне поля.....	27
2.5.2. Связь в дальней зоне поля.....	33
2.6. Влияние окружающей среды.....	43
2.6.1. Потери.....	43
2.6.2. Влияние соседних антенн.....	46
2.6.3. Температура и влажность.....	47
2.7. Резюме.....	47
3. Связь	48
3.1. Введение.....	48
3.2. Процесс связи.....	48
3.3. Сигналы и спектры.....	48
3.3.1. Усреднение по времени, мощность, СКО.....	49
3.3.2. Преобразование Фурье и спектры.....	50
3.4. Кодирование в линии передачи сигнала.....	57
3.4.1. Спектральная плотность мощности сигналов.....	60
3.5. Модуляция.....	69
3.5.1. Кодирование и модуляция в прямой линии.....	70
3.5.2. Кодирование и модуляция в обратной линии.....	73
3.5.3. Модуляция с распределенным спектром.....	74
3.6. Вероятность ошибки.....	75

3.6.1. Импульсные помехи.....	76
3.6.2. Гауссовский шум.....	77
3.7. Обнаружение и коррекция ошибок.....	78
3.8. Резюме.....	79
4. Электромагнитная совместимость и регламенты.....	81
4.1. Введение.....	81
4.2. Необходимость распределения частот.....	81
4.3. Всемирная структура регулирующих организаций.....	82
4.3.1. Организация ИТУ.....	82
4.3.2. Региональные и национальные органы.....	83
4.4. Диапазоны частот RFID.....	84
4.4.1. Диапазоны ISM.....	84
4.4.2. Диапазоны (LF, HF и UHF).....	85
4.5. Напряженность поля, ширина спектра и методы измерения.....	86
4.5.1. Преобразования напряженности поля и мощности.....	86
4.5.2. Ширина спектра.....	88
4.5.3. Детектирование.....	88
4.6. Частоты и ограничения.....	90
4.6.1. Частоты RFID в ближней зоне.....	91
4.6.2. Частоты RFID в дальней зоне.....	97
4.7. Резюме.....	104
5. Аппаратурное исполнение.....	105
5.1. Введение.....	105
5.2. Антенны.....	106
5.2.1. Ближняя зона.....	106
5.2.2. Дальняя зона.....	108
5.3. Интегральные схемы.....	110
5.3.1. Общая схема построения.....	111
5.3.2. Компоненты.....	111
5.3.3. Энергопотребление интегральных схем.....	118
5.4. Соединительные цепи и корпусирование.....	120
5.5. Резюме.....	120
6. Протоколы команд.....	122
6.1. Введение.....	122
6.2. Обзор антиколлизийных алгоритмов.....	122
6.2.1. Анализ антиколлизийных алгоритмов.....	127
6.2.2. Локализация.....	132
6.2.3. Системные вопросы.....	133
6.3. Команды.....	133
6.3.1. Команды и ширина спектра.....	134
6.4. Резюме.....	134
7. Влияние ограничений на технические характеристики.....	135
7.1. Введение.....	135

7.2. Дальность.....	135
7.2.1. Напряженность поля, ориентация и окружающая среда.....	136
7.2.2. Прием, распределение и потребление мощности сигнала.....	137
7.3. Быстродействие и надежность связи.....	140
7.3.1. Скорость передачи данных.....	141
7.3.2. Быстродействие идентификации.....	143
7.4. Аппаратная совместимость и стандартизация.....	144
7.5. Резюме.....	146
8. Заключение.....	147
8.1. Заключение.....	147
8.2. Направления будущих работ и исследований.....	148
Список литературы.....	150

Приложение 1. С.В. Корнеев. К вопросу о дальности действия систем RFID.....

1. Энергетическое и модуляционное уравнения дальности.....	157
2. Зависимость дальности действия от несущей частоты.....	159
3. Зависимость дальности действия от глубины модуляции.....	160
4. Определение технических параметров аппаратуры.....	162
Список литературы.....	163

Приложение 2. И. Девиль, Ж. Дамур, Н. Чаркани. Системы RFID с использованием нейронных сетей.....

1. Введение.....	165
2. Предлагаемая система.....	167
3. Проблема слепого разделения источников и ее решение.....	168
4. Экспериментальные результаты.....	171
5. Обсуждение и заключение.....	182
Список литературы.....	185

Приложение 3. А.С. Гулария, С.В. Корнеев. Стандартизация технологии RFID.....

1. Актуальность внедрения технологии.....	187
2. Эффект от внедрения систем RFID.....	189
3. Проблемы внедрения технологии.....	190
4. Стандартизация систем RFID.....	191
Список литературы.....	193
Послесловие редактора.....	194
Предметный указатель.....	196

Предисловие

Основой данной книги является перевод работы (Т.А. Scharfeld «An Analysis of the Fundamental Constrains on Low Cost Passive Radio Frequency Identification System Design»), представленной в 2001 г. аспирантом Массачузетского Технологического Института Т. Шарфельдом в рамках выполнения требований на соискание степени Master of Science с дополнительными материалами (Приложениями) И. Девиля, Ж. Дамура, Н. Чаркани, С. Корнеева и А. Гуларии.

Радиочастотная идентификация является новой, молодой технологией. Характерным моментом развития этой технологии в настоящее время является практически полное отсутствие научно-технической литературы, посвященной как теоретическим, так и практическим вопросам. Многочисленные статьи в отечественных периодических изданиях носят поверхностный, обзорный характер и, зачастую, содержат существенные погрешности и неточности. Даже за рубежом имеется единственная серьезная монография (Klaus Finkenzeller, RFID Handbook, John Wiley and Sons, Inc., New York, 1999), в которой специалисты могут найти некоторые ответы на вопросы проектирования и практического использования аппаратуры радиочастотной идентификации.

В этой связи работа Т. Шарфельда чрезвычайно актуальна и своевременна. В то же время она далеко не свободна от недостатков технического и стилистического характера, которые, в ряде случаев, потребовали существенной переработки оригинала в процессе перевода. Работа отличается широким и всесторонним рассмотрением проблем радиочастотной идентификации и, одновременно, большой неравномерностью глубины рассмотрения вопросов. Материалы, приведенные в Приложениях, в некоторой степени позволяют дополнить работу Т.А. Шарфельда. Книга несомненно будет уникальным и ценным пособием для студентов, аспирантов и специалистов – системных интеграторов и разработчиков, желающих совершенствовать свой опыт в области перспективной технологии радиочастотной идентификации.

Хочется выразить глубокую благодарность администрации MIT и И. Девилю, любезно предоставивших права на перевод и опубликование авторских материалов, а также сотрудникам Центрального научно-исследовательского радиотехнического института и Научно-технического центра «Альфа-1», бывших первыми читателями и конструктивными критиками русского перевода: М.А. Лякину, Б.М. Чубакову и многим другим, давшим ценные замечания и оказавшим техническую помощь в процессе редактирования и подготовки материалов книги к изданию.

Переводчик, научный редактор С.В. Корнеев

Глава 1

Введение

1.1. Вводные замечания

Использование сети Интернет позволяет осуществить постоянный доступ к информации от самых разнообразных источников. Интернет является средством, при помощи которого только доставляется информация. Во многих случаях информация описывает состояние реальных физических объектов и событий. В других случаях информация существует в виртуальной форме. Виртуальная информация достаточно легко может использоваться и быть доступной людям или механизмам. Серьезной проблемой является преобразование информации из одной формы в другую, из физической в виртуальную. Типичный способ решения этой проблемы предполагает применение ручного труда – человек должен наблюдать и записывать информацию. Такой способ преобразования информации не только неэффективен, но и приводит к ошибкам, неточностям. Известные современные способы, использующие сложные интеллектуальные устройства и обладающие "зрением и чувствами", также позволяют и наблюдать, и записывать, то есть преобразовывать информацию из одной формы в другую. Однако для сбора больших объемов информации такие решения сложны и дороги или требуют выполнения определенных условий и ограничений.

Одним из потенциально возможных решений проблемы преобразования информации является использование технологии радиочастотной идентификации (RFID, в англоязычной литературе используется аббревиатура RFID – **R**adio **F**requency **I**Dentification). При этом информация в физической форме получается за счет бесконтактного считывания данных, которые записаны, в транспондерах или метках, прикрепленных к живым или неживым объектам. Одна из разновидностей систем RFID - пассивные системы - обеспечивает функционирование меток без встроенных источников питания. Потенциально пассивные системы RFID способны иметь чрезвычайно низкую стоимость и тем самым частично обеспечить решение проблемы автоматического снятия физически встроенной информации. Тем не менее, пассивные системы RFID имеют определенные фундаментальные ограничения. Одной из задач данной работы является рассмотрение этих ограничений.

©Т.А. Scharfeld, 2001.

©С.В. Корнеев, перевод на русский язык и научная редакция, 2004.

Для лучшего понимания этого вопроса целесообразно рассмотреть цепь доставки товаров. Известно, что обладание информацией позволяет повысить эффективность и осуществить огромное сокращение расходов производителям, дистрибьюторам, продавцам и любым другим звеньям цепи доставки товаров потребителю [1]. При помощи сети Интернет решается проблема только доставки информации, но не решается проблема собственно ее извлечения или получения. Современные технические решения предполагают ручной ввод информации или ее ручное сканирование при помощи штрих-кодов. Более совершенные системы обеспечивают автоматическое сканирование штрих-кодов или содержат интеллектуальные визуальные считывающие устройства. Решения с использованием ручного труда достаточно дороги, обладают невысоким быстродействием и подвержены ошибкам. Используемые в настоящее время решения, основанные на применении средств автоматизации, также могут быть дорогостоящими, сложными и часто имеют эксплуатационные ограничения, обусловленные влиянием окружающей среды. При этом использование технологии RFID является потенциальным решением проблемы преобразования информации из физической в виртуальную форму.

Системы RFID позволяют считывать информацию, находящуюся вне пределов видимости. Идентификационный код хранится в метке, состоящей из микрочипа, прикрепленного к антенне. Приемопередатчик, часто называемый интеррогатором или считывателем, имеет связь с меткой с помощью телекоммуникации. Сами по себе системы RFID не пригодны для использования. Хотя они и могут автоматически считывать коды, эти коды должны быть заложенными в базу данных и соответствовать каким-либо объектам. Должна быть создана необходимая сетевая инфраструктура или иная система, способная хранить и обрабатывать идентификационные коды, поддерживать сбор и доставку информации, которая получена системами RFID [2]. Такая сетевая инфраструктура должна эффективно обеспечивать хранение и доставку огромного количества информации.

Целью данной работы является рассмотрение и анализ всех элементов информационной системы, включая метки, считыватели, системы сбора информации (host systems), базы данных и сетевое окружение.

Интернет и связанная с ним инфраструктура в настоящее время является функционирующей и быстро развивающейся технологией. Технология RFID, хотя исторически и старше Интернета, относительно молода и все еще недостаточно опробована для надежного, большого по объему и низкого по стоимости использования. Тем не менее, в настоящее время техническая реализация аппаратуры RFID уже удовлетворительна и постоянно совершенствуется. Все же стоимость ее еще относительно высока (минимально 0,5 доллара за метку и более 100 долларов за считыватель), но в ближайшем будущем с улучшением технологии будет быстро снижаться.

Технологические инновации и рост объемов использования улучшают техническую реализацию аппаратуры и снижают ее стоимость. Грамотная стандартизация приведет к росту объемов использования и также снизит стоимость аппаратуры RFID. Для более подробного рассмотрения и направлений развития технологии RFID обратимся к истории ее возникновения.

1.2. Краткий исторический обзор

RFID является простым сокращенным названием, данным семейству технологических датчиков, которые появились в последние 50 лет. Первое широкое использование технологии RFID произошло в течение Второй мировой войны [3]. Британские вооруженные силы на своей авиации и авиации союзников использовали транспондеры, способные при поступлении запрашивающего сигнала отвечать соответствующим идентификационным сигналом. Эта технология не позволяла реализовать точную идентификацию объекта, но была достаточной для определения своих самолетов. Эта технология, названная идентификацией "свой - чужой" (IFF), нашла дальнейшее развитие, и ее последующие поколения сейчас используются как в военной, так и в гражданской авиации.

После начального использования во Второй мировой войне технология RFID получила существенное развитие. В 60-х и 70-х годах из-за необходимости безопасного и секретного сопровождения военного персонала и оборудования лаборатории различных государств разработали идентификационные технологии. В конце 70-х годов из научных лабораторий в Лос Аламосе выделились две компании, которые коммерциализировали эту технологию. Первые применения RFID были предназначены для идентификации и определения температуры рогатого скота [4]. В начале 80-х годов железнодорожные компании стали использовать эту технологию для идентификации и сопровождения вагонов.

В этих применениях обычно использовался UHF диапазон (900/1800 МГц). В течение 80-х годов некоторые компании США и Европы начали разрабатывать технологии для работы на других частотах, с другими источниками мощности, объемами памяти и рядом других функций. Во второй половине 80-х, когда подключились крупные компании электронной промышленности, начался существенный технический прогресс и снижение размеров и стоимости аппаратуры.

В конце 80-х и в течение 90-х годов по мере совершенствования аппаратуры, снижения ее размеров и стоимости, появились новые направления использования RFID. Некоторые из них: автоматические платежные системы на автострадах, системы контроля безопасности и

доступа, автомобильные иммобилайзеры, системы контроля багажа в авиaperевозках, инвентаризация товаров и имущества и смарт-карты [5–8].

Для дальнейшего расширения спектра применений RFID необходимо обеспечить высокий технический уровень исполнения, малые габариты и низкую стоимость, совершенствование уровня разработки и производства аппаратуры. В настоящее время проводится большая работа и возникает большое число компаний, занимающихся инновациями, снижением размеров, цены и повышением технического уровня в этой области. Возникновение Интернета и развитие информационных технологий позволяет найти новые возможности применения технологии RFID.

Однако для широкомасштабного глобального внедрения технологии RFID необходимым является не только техническое совершенствование аппаратуры, снижение ее габаритов и цены – системы разных продавцов должны быть совместимыми друг с другом и способными работать в условиях различных административных регламентов, как местных, так и международных. В настоящее время продолжается работа по созданию стандартов для различных применений технологии RFID. Стандарты, однако, с одной стороны, могут способствовать техническому прогрессу, а с другой – подавлять конкуренцию, инновации и стремление к совершенствованию аппаратуры. В дальнейшем мы рассмотрим основные ограничения, которые налагает стандартизация на развитие RFID.

1.3. Классификация систем RFID

В процессе развития технологии RFID возникло несколько типов систем. Они могут классифицироваться несколькими способами. Термин RFID включает довольно широкий класс идентификационных устройств.

Все системы RFID содержат считыватели и метки. Считыватели извлекают информацию, которая хранится или собирается меткой. Считыватели размещаются в некоторой точке пространства, а метки прикрепляются к объектам. Из-за того, что метки устанавливаются на различные объекты, на их характеристики налагаются существенные ограничения по исполнению, размерам и стоимости. С этими характеристиками связаны различные классификации систем RFID.

Одна широко распространенная классификация разделяет метки на "чиповые" и "бесчиповые". Чиповые метки содержат интегральную микросхему – чип, а бесчиповые ее не содержат. Другая классификация, которая подразделяет чиповые метки, выделяет пассивные, полуактивные и активные метки. Пассивные метки не содержат ни элемента питания, ни активного передатчика; полуактивные метки содержат элемент питания, но не имеют активного передатчика; активные метки содержат и то и другое.

Еще одна классификация подразделяет метки на только считываемые (read only) и считываемые/записывающие (read/write). Только считываемые метки имеют или только считываемую память, или память, которая однократно программируется и многократно считывается. Считываемые/записывающие метки позволяют однократно записывать и многократно перезаписывать информацию.

В дальнейшем основное внимание будет уделено пассивным чиповым только считываемым меткам. Кратко рассмотрим классификацию меток и объясним, почему на нее обращается внимание.

1.3.1. Чиповые и бесчиповые метки

Так как наше внимание обращается на метки самой низкой стоимости, которые обеспечивают минимум функциональных возможностей – простые только считываемые устройства с постоянным уникальным идентификационным кодом – может показаться, что бесчиповые метки могли бы быть оптимальными. Тем не менее, в дальнейшем мы в основном будем рассматривать чиповые метки по двум причинам:

- 1) метка должна обладать достаточной памятью, чтобы хранить уникальный идентификационный номер большого числа идентифицируемых объектов;
- 2) считыватель должен обладать возможностью считывания множества меток, находящихся в его зоне действия.

Для того чтобы однозначно идентифицировать множество производимых объектов, схема памяти должна иметь возможность хранения достаточного числа уникальных кодов. В ближайшем будущем достаточно 64 бита информации, однако в перспективе оптимальным считается 96 бит [9]. Большинство бесчиповых меток в настоящее время позволяют хранить 24 бита или меньше, хотя некоторые позволяют хранить 64 бита [10]. Это, однако, приводит к повышению их стоимости.

Из-за возрастания числа и снижения размера объектов, на которые устанавливаются метки, необходимо, чтобы считыватель был способен одновременно считывать множество меток, находящихся в зоне его действия; причем метки могут размещаться близко друг от друга. В настоящее время наилучшим способом решения такой задачи – коллизии сигналов – является наделение самих меток некоторым интеллектом. Методы пространственного выделения одной метки среди многих других, что приводит к разрешению коллизии, хотя и потенциально возможны, все еще трудно реализуемы. В главе 7 мы рассмотрим способы идентификации множества меток.

Хотя бесчиповые метки и анонсируют достижение больших объемов памяти и улучшение антиколлизийных функций, а их схемы, в будущем,

смогут изготавливаться методом непосредственной печати на некремниевые подложки, в ближайшем будущем практически только чиповые метки могут удовлетворить большинство применений, требующих идентификации и информационного сопровождения товаров.

1.3.2. Пассивные, полуактивные и активные метки

Различие между пассивными, полуактивными и активными метками состоит в наличии источника питания и передатчика. Пассивные метки не содержат ни источника питания, ни передатчика. Полуактивные метки содержат источник питания, но не содержат передатчика. Активные метки содержат как источник питания, так и передатчик.

Активные метки имеют наилучшие характеристики. Дальность может достигать километров, а связь со считывателем надежная и быстрая. Однако наличие источника питания и передатчика приводит к высокой стоимости.

Полуактивные метки по сравнению с пассивными имеют более высокую дальность (до нескольких десятков метров) и из-за этого могут иметь достаточно высокие функциональные возможности. Однако это также приводит к повышению их стоимости.

Пассивные метки обладают дальностью менее 10 метров и более зависимы от регламентных ограничений и влияния окружающей среды. Тем не менее они имеют максимальный рыночный потенциал из-за наименьшей стоимости.

1.3.3. ROM и WORM метки

Любые чиповые метки могут быть только считываемыми или считываемыми/записываемыми. Пассивные метки, как правило, бывают только считываемыми. Только считываемые метки программируются идентификационным кодом в процессе производства или при установке на определенный объект. Память таких меток может быть или только читаемой памятью (ROM) или однократно программируемой и многократно читаемой (WORM). В дальнейшем в главе 5 мы рассмотрим такие разновидности памяти.

Считываемые/записываемые метки могут многократно перепрограммироваться в процессе их эксплуатации. Обычно они имеют идентификационный код или серийный номер, который записывается в процессе производства. Также в них может записываться разнообразная дополнительная информация. Такие метки многофункциональны, однако это приводит к возрастанию их стоимости.

Из-за низкой стоимости наибольшим рыночным потенциалом обладают только считываемые метки, а в совокупности с хорошо спроектированной распределенной базой данных они обеспечивают практически такие же функциональные возможности, как и считываемые/записывающие метки.

1.4. Предположения

Перед рассмотрением технических характеристик аппаратуры RFID еще раз обратимся к принятым предположениям. Будем считать, что метки являются пассивными и только считываемыми. Предположим, что дальность считывания меток не превышает 6 метров, а в системе присутствует один или несколько считывателей. Также предположим, что существует сетевая инфраструктура, посредством которой уникальные идентификационные коды, соответствующие определенным объектам, могут быть привязаны к этим объектам [2]. При этом идентификационные коды разделены на логические поля, в которые занесены код товара, код его владельца, код производителя и серийный номер метки [9].

1.5. Характеристики аппаратуры

История развития показывает, что применения RFID изменялись в соответствии с развитием собственно технологических возможностей. Каждое из применений налагает свои специфические требования к аппаратуре. Для того чтобы достаточно подробно рассматривать особенности технологии RFID, как для лучшего понимания ее функциональных возможностей, так и для лучшего проведения разработки аппаратуры, необходимо оценивать ее функции на основе технических характеристик. Принимая во внимание большое разнообразие применений технологии RFID, это сложная задача. Тем не менее, существуют основные характеристики, общие для всех применений. Такими характеристиками является стоимость, габариты и технические параметры аппаратуры. В свою очередь основными техническими параметрами аппаратуры RFID являются дальность считывания, быстродействие, надежность связи и совместимость систем различных поставщиков с учетом административных норм в различных регионах мира.

Наиболее жесткими и высоко требовательными являются применения, относящиеся к объектам низкой стоимости и к объектам, находящимся непосредственно в распоряжении потребителей. Двумя типичными примерами является идентификация почтовых отправлений [11] и товаров, находящихся в розничной продаже. Считыватели должны обеспечивать быстрое считывание большого числа меток на расстоянии нескольких

метров в относительно сложных условиях функционирования аппаратуры в окружающем ее пространстве.

Чрезвычайно важной характеристикой аппаратуры RFID является ее стоимость. Так как маркированные предметы находятся в розничной продаже, а количество этих предметов огромно, совсем небольшое увеличение стоимости одной метки приводит к очень большому возрастанию совокупной цены аппаратуры. В зависимости от габаритов объекта также чрезвычайно важны размеры метки. Кратко рассмотрим технические параметры, цену и габариты аппаратуры RFID.

1.5.1. Технические параметры

Основными техническими параметрами аппаратуры RFID являются: дальность, быстродействие, надежность связи и совместимость. Определение конкретных значений технических параметров требует достаточно точной оценки условий применения аппаратуры.

Широкая область применения – идентификация объектов, находящихся в розничном обращении, может быть подразделена на несколько практически самостоятельных применений, каждое из которых имеет различные конфигурации. Так, например, мы можем самостоятельно рассматривать каждое звено цепи поставок: предприятие - изготовитель, дистрибьюторский центр, магазины продажи, дома потребителей и центры утилизации отходов, а также транспортировку между ними.

На предприятии – изготовителе применениями являются сборка, сортировка и отгрузка товаров. На дистрибьюторских центрах применениями могут быть: прием, сортировка, подбор (picking) и отгрузка. В магазинах продажи применениями могут быть: получение, сортировка и другие операции, требующие инвентаризации товаров. Домашние применения могут заключаться в инвентаризации, а применения в центрах утилизации отходов – в сортировке. Логистические операции требуют иных применений.

Вполне вероятно, что если товары розничной продажи маркированы, то должны также быть маркированы коробки для них, поддоны для коробок и контейнеры, в которых находятся поддоны. Метки для поддонов и контейнеров вероятно должны допускать повторное использование, поэтому ценовые ограничения здесь не слишком строги. Коробок по количеству меньше, чем предметов в них, поэтому ценовые ограничения для них менее строги, чем в случае предметов, но более строги, чем для поддонов.

Предположим, что содержимое контейнеров, поддонов и коробок, которые индивидуально промаркированы, должно быть известным. Идентификация предметов розничной продажи необходима в местах, где

они не упакованы. Такая ситуация возможна на предприятиях и дистрибьюторских центрах, однако она наиболее вероятна в магазинах, в домах и в центрах утилизации отходов. Особенно требовательны применения при операциях проверки в магазинах. Считыватели и метки должны справляться с возможными воздействиями расположенных неподалеку других считывателей, с отражениями и влияниями от металлических предметов и людей, находящихся в зоне действия аппаратуры. Может потребоваться точная и быстрая идентификация большого числа случайно расположенных предметов. Часто желательна максимальная дальность считывания в несколько метров. Точные технические параметры в любом случае определяются конкретными реальными условиями применения. Считыватели могут устанавливаться на воротах и порталах, на подвижных тележках или контейнерах, или быть позиционированными в нужной зоне или области. Зона действия может быть частично ограниченной, либо совсем неограниченной – может потребоваться отделение одного предмета от других предметов и его индивидуальное считывание или считывание в группе других предметов, движущихся на достаточно большой пешеходной скорости.

1.5.2. Стоимость и размеры

С точки зрения конечного пользователя, вне всякого сомнения, стоимость должна быть минимальной. Положим, что метками маркированы товары розничной продажи. Небольшое снижение стоимости одной метки приведет к огромным сбережениям конечного пользователя. В настоящее время стоимость метки составляет 5 центов и менее при очень больших объемах производства. Минимизация стоимости при таких масштабах ставит серьезнейшие проблемы при разработке. Каждое звено системы должно быть оптимизировано, во-первых, для снижения стоимости и, во-вторых, для достижения необходимых характеристик – не только метки с их электронными схемами и антеннами, но также вся технология производства, сборки и использования конечной продукции. При проектировании должны творчески учитываться ограничения, наложенные на систему.

Размеры также, конечно, исключительно важны. Метки, как минимум, должны быть меньше маркируемого объекта. Так как уже производятся микрочипы для меток размером менее 0,5 кв. мм, антенны всегда будут ограничивающим фактором. Поскольку технические параметры, особенно дальность, зависят от размеров и формы антенн, а размеры антенны существенно зависят от рабочей частоты, метки на очень маленьких объектах не будут обладать такими техническими характеристиками, как метки с большими антеннами, более соответствующими рабочему диапазону частот. На 5,8 ГГц, например, оптимальный размер антенны равен 2,5 см, в то время как на частоте 915 МГц – 16 см. Антенны с

размерами меньшими оптимального могут иметь удовлетворительные параметры, но все-таки хуже, чем у оптимальных антенн. Простое повышение рабочей частоты не всегда является выходом, так как появляются другие ограничения, такие как распространение радиоволн, затухание и административные нормы. Размеры в очень большой степени зависят от конкретного применения и конфигурации системы. В некоторых случаях могут потребоваться весьма небольшие расстояния считывания. Тогда и размеры могут быть очень малыми.

Все вышесказанное по поводу размеров и стоимости оказывается тесно связанным с техническими характеристиками, такими как дальность, быстродействие и надежность связи. Все эти параметры в комбинации с основными ограничениями являются основными исходными данными при проектировании систем RFID.

1.6. Функции, ограничения и структура работы

После того как мы оговорили предположения и определили важнейшие параметры системы RFID, необходимо рассмотреть основные ограничения. Так как на этих ограничениях базируется структура данной работы, мы рассмотрим структуру и ограничения одновременно. Прежде всего, однако, будет полезно еще раз рассмотреть компоненты системы и ее функции.

1.6.1. Компоненты и функции системы

В общем случае, как показано на рис.1.1, система RFID состоит из четырех основных составляющих: система сбора информации (хост), считыватель, множество меток и канал распространения, посредством которого взаимодействуют считыватель и метки.

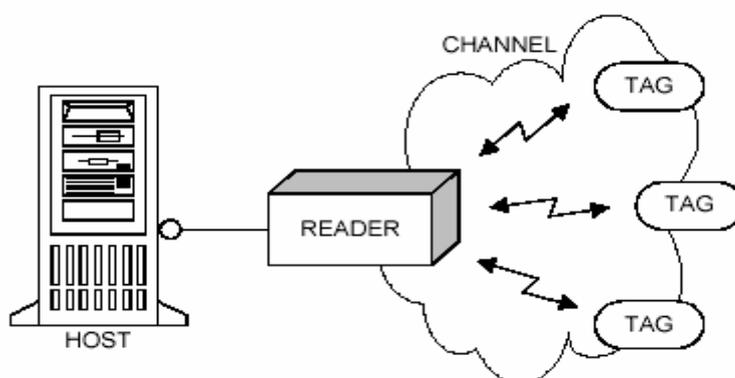


Рис 1.1. Четыре основных компонента системы RFID: хост, считыватель, канал распространения и метки.

(host – хост, - считыватель, channel - канал распространения, tag - метка)

Мы рассмотрим только те компоненты системы, которые непосредственно связаны с меткой, а именно: высокочастотный интерфейс считывателя (радиоинтерфейс), собственно метки и канал распространения сигнала. Функции, выполняемые этими элементами с их характеристиками, и ограничения, которые определяют проектирование элементов, показаны на рис. 1.2.

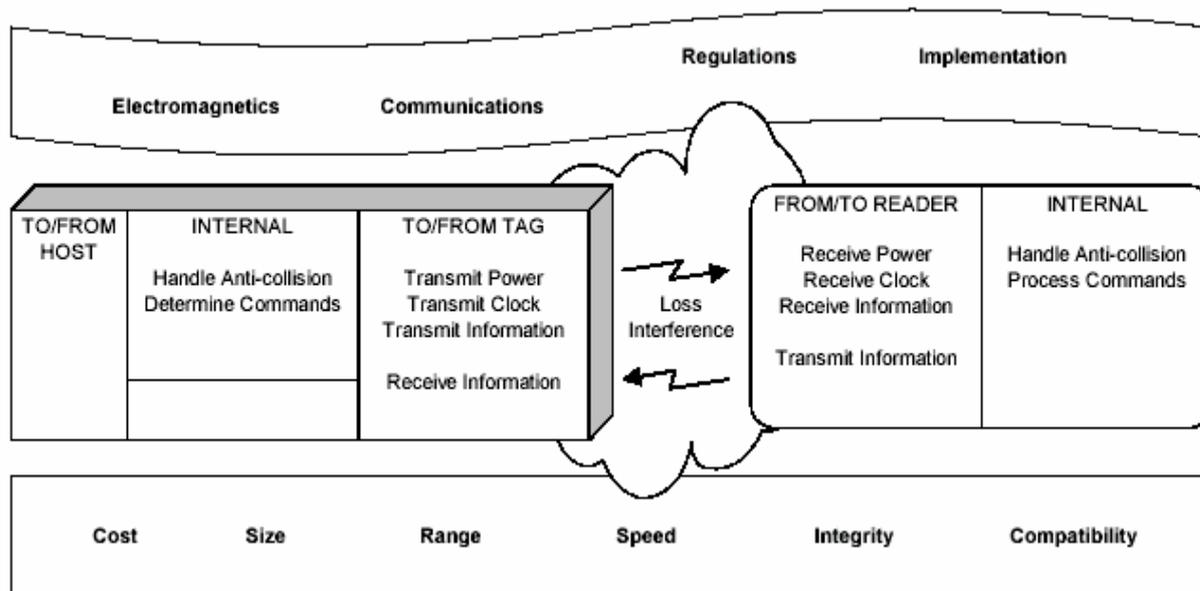


Рис. 1.2. Функции считывателя, канала и метки и влияющие на них характеристики и ограничения.

(Electromagnetics - электродинамика, Communications - связь, Regulations - регламенты, Implementation - техническое исполнение, To/From Host - К/От хоста, Internal - Внутренние функции, Handle Anti-collision - Анतिकоллизионные алгоритмы, Determine Commands - Командные протоколы, To/From Tag - К/От метки, Transmit Power - Излучаемая мощность, Transmit Clock - Передаваемая синхронизация, Transmit Information - Передаваемая информация, Receive Information - Принимаемая информация, Loss - Потери, Interference - Интерференция, To/From Reader - К/От считывателя, Receive Power - Принимаемая мощность, Receive Clock - Принимаемая синхронизация, Process Commands - Обработка командных протоколов, Cost - Цена, Size - Стоимость, Range - Дальность, Speed - Быстродействие, Integrity - Надежность, Compatibility - Совместимость.)

Из рис.1.2 видно, что функции считывателя и метки взаимно дополняют друг друга. Внутренние функции, как метки, так и считывателя, включают алгоритмы верхнего уровня и командные протоколы, необходимые для идентификации одной или множества меток, находящихся в зоне действия считывателя. Команды в общем случае основаны на алгоритмах, которые называются антиколлизионными алгоритмами, так как они реализуют снижение вероятности ошибок от совпадений ответов меток,

поступающих по запросу считывателя. Интерфейсные операции содержат команды нижнего уровня, необходимые для реализации задач команд верхнего уровня. Считыватель должен обеспечивать излучение мощности и передачу информационных и синхронизирующих сигналов. Информационные сигналы необходимы для обеспечения идентификации множества меток и реализации антиколлизийных алгоритмов. Синхронизация необходима для обеспечения работы цифровой части электронной схемы метки. В зависимости от несущей частоты синхросигналы генерируются непосредственно из высокочастотного сигнала или при помощи модуляции его несущей.

Метка обеспечивает прием энергии, а также обработку информационных и синхронизирующих сигналов, передаваемых считывателем. После внутренней высокоуровневой обработки принятой информации метка может передать считывателю информацию в виде полного идентификационного кода или его части. В пассивных системах RFID процесс передачи информации осуществляется за счет модуляции сигнала считывателя. Считыватель, в свою очередь, должен принять информацию и передать ее обратно на внутренний верхний уровень функционирования системы.

1.6.2. Организация и структура работы

Проектирование и практическая реализация вышеперечисленных функций определяется как характеристиками системы, так и основными ограничениями. Ограничения связаны с особенностями электродинамики и связи, которые, определяют способы взаимодействия считывателя и метки, а также особенности аппаратных и административных ограничений, которые, в свою очередь, тоже налагают ограничения на эти способы.

В главе 2 мы рассмотрим электромагнитные поля и волны, которые связывают метку и считыватель. Мы покажем, как электромагнитные поля и волны создаются антеннами и как параметры и поведение этих волн зависят от несущей частоты и расстояния от антенны. Мы рассмотрим, как системы RFID используют параметры и характеристики волн для обеспечения передачи мощности и информации. В заключительной части главы мы оценим влияние различных условий окружающей среды на поведение полей и волн и на взаимосвязь между считывателем и меткой.

В главе 3 мы рассмотрим процесс связи между считывателем и меткой, в том числе и теорию, которая его описывает. Мы рассмотрим и проанализируем характеристики различных кодирующих и модулирующих схем, обычно используемых в технологии RFID. Также мы обсудим методы обнаружения сигналов и вопросы, связанные с надежностью связи и обнаружением ошибок.

В главе 4 мы рассмотрим регламенты (административные нормы), которые определяются как электромагнитными, так и коммуникационными (связными) аспектами. Мы будем рассматривать регламенты в трех регионах мира: в Европе, в США и в Японии. Также мы обсудим различные методы введения регламентов, и какое влияние это оказывает на технологию RFID.

В главе 5 мы рассмотрим ограничения, связанные с физическим исполнением аппаратуры. Также обсудим характеристики антенн, которые влияют на их технические параметры, размеры и стоимость. Кроме того здесь мы рассмотрим микросхемы и как проектирование микросхем влияет на их параметры и стоимость.

В главе 6 мы рассмотрим и проанализируем антиколлизийные алгоритмы и командные протоколы, необходимые для идентификации множества меток в поле действия считывателя.

В главе 7 мы соотнесем фундаментальные ограничения с важнейшими техническими характеристиками и рассмотрим их взаимное влияние.

И, наконец, в главе 8 мы сделаем заключение и попытаемся представить направления будущих работ в области технологии RFID.

Глава 2

Электродинамика и антенны

2.1. Введение

В системах RFID считыватели взаимодействуют с метками при помощи распространяющихся в пространстве электромагнитных полей и волн. Для того чтобы объяснить работу и грамотно проектировать эти системы, необходимо понять, как создаются эти поля и волны, как они модулируются и принимаются. В этом разделе мы проведем обзор принципов и поведения электромагнитных полей и волн. Необходимо понять:

- как создаются поля, какова их протяженность, доступная мощность, их изменение с расстоянием, углами, ориентацией и поляризацией, а также как они излучаются и принимаются;
- каким образом типы антенн, их размеры и форма влияют на свойства поля;
- каковы возможности максимизации принимаемой мощности и информации, применяя настройки и согласование;
- как изменяются реактивные и излучающие свойства линии связи между считывателем и меткой в зависимости от свойств канала распространения сигнала и его окружающей среды.

Все это поможет нам понять физические ограничения систем RFID и основы управления их функциями. Кроме того, это поможет нам понять другие фундаментальные ограничения, включая теорию связи и административные регламенты. В итоге мы обсудим, как эти различные факторы влияют на исполнение и стоимость систем RFID.

2.2. Уравнения Максвелла и основы электродинамики

Когда статическая электрическая линия поля, создаваемого электрическим зарядом, распространяется в некотором направлении, излучается электрическое поле. Когда этот электрический заряд осциллирует, он излучает осциллирующее электрическое поле. Вариация электрического поля в пространстве и времени вызывает связанное с ним магнитное поле. Уравнения Максвелла описывают поведение электромагнитных волн в любой точке пространства и в любой момент времени относительно расположения и движения заряженных частиц [12]. Мы можем представить эти уравнения в гармонической (синусоидальной) форме с частотой ω в виде соотношения $\partial/\partial t = -j\omega t$:

$$\nabla \times \mathbf{E} = j\omega \mathbf{B}, \quad (2.1)$$

$$\nabla \times \mathbf{H} = \mathbf{J} - j\omega \mathbf{D}, \quad (2.2)$$

$$\nabla \cdot \mathbf{D} = \rho, \quad (2.3)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (2.4)$$

где \mathbf{E} – напряженность электрического поля (В/м),

\mathbf{H} – напряженность магнитного поля (А/м),

\mathbf{B} – магнитная индукция (Т),

\mathbf{D} – электрическая индукция (К/м²),

\mathbf{J} – плотность электрического тока (А/м²),

ρ – плотность электрического заряда (К/м³).

Уравнение неразрывности электрического заряда и тока записывается следующим образом:

$$j\omega\rho + \nabla \cdot \mathbf{J} = 0. \quad (2.5)$$

Совокупность уравнений Максвелла (2.1) – (2.4) с уравнением (2.5) представляет собой фундаментальные уравнения электродинамики.

Усредненная по времени мощность, переносимая электрическим и магнитным полями через единицу поверхности, определяется вектором Пойнтинга:

$$\mathbf{S} = \frac{1}{2} \operatorname{Re} \{ \mathbf{E} \times \mathbf{H}^* \}, \quad (2.6)$$

где \mathbf{H}^* – вектор, комплексно сопряженный вектору \mathbf{H} , а \mathbf{S} – плотность потока мощности, имеющая размерность Вт/м².

В области пространства, не имеющей зарядов и источников тока, связанное с магнитным полем и ортогональное ему электрическое поле, распространяется в окружающей среде в направлении, перпендикулярном обоим полям. Это и есть электромагнитная волна. Антенна является устройством, предназначенным для излучения электромагнитных полей и волн.

В дальнейшем, для описания полей и волн мы будем использовать несколько параметров. Эти параметры могут быть заданы в общей форме для описания взаимодействия с любыми материалами. Мы рассмотрим параметры среды без потерь и среды с потерями. Волновое число в свободном пространстве:

$$k_0 = \omega (\mu_0 \varepsilon_0)^{1/2} = \omega / c = 2\pi / \lambda_0 = \beta, \quad (2.7)$$

где c – скорость света $\approx 3 \times 10^8$ (м/с), ϵ_0 – диэлектрическая проницаемость свободного пространства $= 8,8542 \times 10^{-12}$ (ф/м), μ_0 – магнитная проницаемость свободного пространства, равная $4\pi \times 10^{-7}$ (Гн/м).

Диэлектрическая проницаемость связывает напряженность электрического поля \mathbf{E} и электрическую индукцию \mathbf{D} соотношением $\mathbf{D} = \epsilon_0 \mathbf{E}$. Магнитная проницаемость связывает напряженность магнитного поля \mathbf{H} и магнитную индукцию \mathbf{B} соотношением $\mathbf{B} = \mu_0 \mathbf{H}$.

Другой полезный параметр – сопротивление свободного пространства:

$$\eta_0 = (\mu_0 / \epsilon_0)^{1/2}. \quad (2.8)$$

Используя фундаментальные уравнения электромагнетизма и введенные параметры, мы сможем понять, как создаются электромагнитные поля и волны.

2.3. Антенны и окружающее пространство

Для лучшего понимания соотношений между электромагнитными полями и волнами полезно рассмотреть, как они формируются антенной. При помощи уравнений Максвелла мы можем вычислить электрическое и магнитное поля, создаваемые и излучаемые антенной.

С целью иллюстрации рассмотрим две антенны, предполагая, что их максимальные размеры значительно меньше длины волны сигнала. Мы рассмотрим идеальный диполь, известный как диполь Герца, и малую рамку (рис. 2.1). Идеальный диполь является бесконечно малым элементом, по которому протекает ток с равномерным по его длине распределением амплитуды и фазы. Малая рамка является замкнутой токовой петлей с периметром, меньшим четверти длины волны сигнала. Малая рамка и идеальный диполь дуальны по своим свойствам.

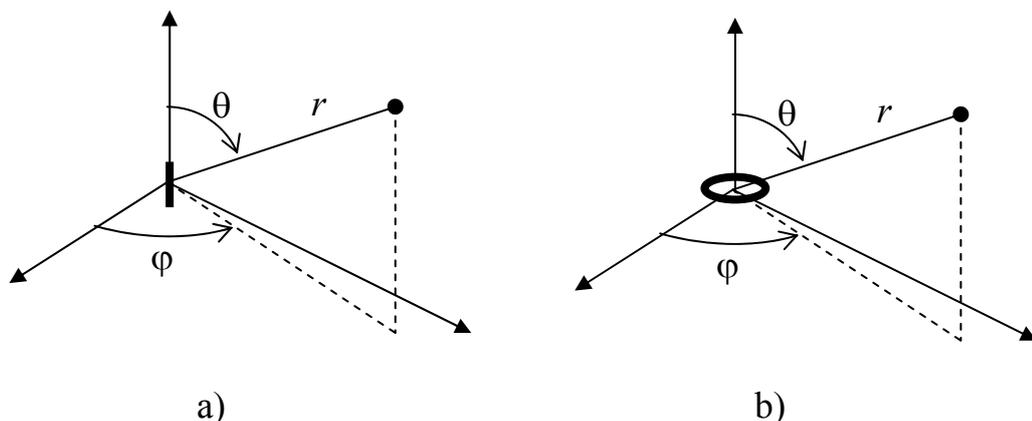


Рис.2.1. а) – идеальный диполь, б) – малая рамка

Общий подход к определению излученного электрического и магнитного полей состоит, прежде всего, в вычислении векторного потенциала, выраженного через плотность тока. Из векторного потенциала определяется электрическое поле, а затем, соответственно, из электрического поля определяется магнитное поле.

Поля, создаваемые идеальным диполем длиной dl , могут быть записаны в виде [14]:

$$\begin{aligned} \mathbf{E} = & -(I dl / 4\pi) \eta_0 \beta^2 2 \cos \theta [(j\beta r)^{-2} + (j\beta r)^{-3}] e^{-j\beta r} \hat{\mathbf{r}} - \\ & -(I dl / 4\pi) \eta_0 \beta^2 \sin \theta [(j\beta r)^{-1} + (j\beta r)^{-2} + (j\beta r)^{-3}] e^{-j\beta r} \hat{\boldsymbol{\theta}}, \end{aligned} \quad (2.9)$$

$$\mathbf{H} = -(I dl / 4\pi) \beta^2 \sin \theta [(j\beta r)^{-1} + (j\beta r)^{-2}] e^{-j\beta r} \hat{\boldsymbol{\phi}}. \quad (2.10)$$

Из рассмотрения этих уравнений видна зависимость электрического и магнитного полей от расстояния r до диполя. Когда $\beta r \ll 1$ (или $r \ll \lambda/2\pi$), в уравнениях (2.9) и (2.10) преобладает третья степень r . На таких расстояниях напряженность электрического поля убывает пропорционально $1/r^3$, а напряженность магнитного поля – пропорционально $1/r^2$. Мы будем называть эту область ближней зоной антенны.

Когда расстояние r значительно больше $\lambda/2\pi$, преобладает первая степень и напряженность обоих полей – электрического и магнитного – убывает пропорционально $1/r$. Будем называть эту область дальней зоной антенны.

Дополнительно рассматривая уравнения, заметим, что в ближней зоне, где $\beta r \ll 1$, величина $e^{-j\beta r}$ приближается к 1. При этом уравнения электрического и магнитного полей упрощаются:

$$\mathbf{E}^{\text{nf}} = j(I dl / 4\pi \beta r^3) \eta_0 (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\boldsymbol{\theta}}), \quad (2.11)$$

$$\mathbf{H}^{\text{nf}} = (I dl / 4\pi r^2) \sin \theta \hat{\boldsymbol{\phi}}. \quad (2.12)$$

Из уравнений (2.11) и (2.12) видно, что электрическое поле является мнимым, то есть сдвинутым относительно магнитного поля на $\lambda/4$ (или на 90 градусов). Это является признаком реактивности поля – энергия остается приблизительно постоянной и перераспределяется между двумя полями. Оценка вектора Пойнтинга показывает, что в этом случае отсутствует

реальный поток энергии. Электрическое и магнитное поля в ближней зоне практически не связаны друг с другом и являются квазистатическими. Дальнейшее упрощение покажет, что в этом случае малый диполь является практически статическим электрическим диполем.

В дальней зоне, где $\beta r \gg 1$, доминирует первая степень r и уравнения (2.9) и (2.10) поля можно упростить:

$$\mathbf{E}^{\text{ff}} = j (I dl / 4\pi r) \eta_0 \beta e^{-j\beta r} \sin\theta \hat{\boldsymbol{\theta}}, \quad (2.13)$$

$$\mathbf{H}^{\text{ff}} = j (I dl / 4\pi r) \beta e^{-j\beta r} \sin\theta \hat{\boldsymbol{\phi}}. \quad (2.14)$$

Отсюда видно, что электрическое и магнитное поля синфазны, а отношение их величин $E_\theta/H_\phi = \eta_0$ – импеданс свободного пространства. Оба поля убывают пропорционально $1/r$. Анализ вектора Пойнтинга показывает, что он является реальной, действительной величиной, что соответствует передаче энергии в пространстве. Совокупность электрического и магнитного поля в дальней зоне представляет собой электромагнитную волну.

Теперь рассмотрим поля, создаваемые малой рамкой:

$$\mathbf{E} = (I dl / 4\pi) \eta_0 \beta^2 \sin\theta [1/(j\beta r) + 1/(j\beta r)^2] e^{-j\beta r} \hat{\boldsymbol{\phi}}, \quad (2.15)$$

$$\begin{aligned} \mathbf{H} = & -(I dl / 4\pi) \beta^2 2\cos\theta [1/(j\beta r)^2 + 1/(j\beta r)^3] e^{-j\beta r} \hat{\mathbf{r}} \\ & -(I dl / 4\pi) \beta^2 \sin\theta [1/(j\beta r) + 1/(j\beta r)^2 + 1/(j\beta r)^3] e^{-j\beta r} \hat{\boldsymbol{\theta}}. \end{aligned} \quad (2.16)$$

Компоненты полей в ближней зоне:

$$\mathbf{E}^{\text{nf}} = -(I dl / 4\pi r^2) \eta_0 \sin\theta \hat{\boldsymbol{\phi}}, \quad (2.17)$$

$$\mathbf{H}^{\text{nf}} = j(I dl / 4\pi \beta r^3) (2\cos\theta \hat{\mathbf{r}} + \sin\theta \hat{\boldsymbol{\theta}}) \quad (2.18)$$

и в дальней зоне:

$$\mathbf{H}^{\text{ff}} = j (I dl / 4\pi r) \beta e^{-j\beta r} \sin\theta \hat{\boldsymbol{\theta}}, \quad (2.19)$$

$$\mathbf{E}^{\text{ff}} = -j(I dl / 4\pi r) \eta_0 \beta e^{-j\beta r} \sin\theta \hat{\boldsymbol{\phi}}. \quad (2.20)$$

Замечая схожесть в уравнениях полей идеального диполя и малой рамки, мы видим, что малая рамка дуальна идеальному диполю.

Часто определяют еще одну область пространства, обычно называемую Рэлеевской областью. В этой области целесообразно аппроксимировать сферическую волну плоской волной. Плоская волна является электромагнитной волной с постоянным направлением, интенсивностью и фазой на бесконечных плоскостях, перпендикулярных направлению распространения. Плоские волны реально не существуют, однако на достаточно большом расстоянии от источника, такая аппроксимация возможна и целесообразна. Эта область задается условием:

$$r > 2D^2 / \lambda, \quad (2.21)$$

где r – расстояние от антенны, а D – протяженность излучающей апертуры антенны. В дальнейшем, когда мы будем рассматривать электромагнитные волны, мы предполагаем, что рассматриваем не только дальнюю зону, но одновременно и Рэлеевскую зону.

В дополнение, чтобы понять соотношения между электромагнитными полями и волнами, мы должны также учитывать, что электромагнитные поля ведут себя совершенно по-разному в ближней и дальней зонах. В ближней зоне поля являются реактивными и квазистатическими, тогда как в дальней зоне они представляют собой распространяющиеся волны. Это особенно важно применительно к системам RFID. Те системы, которые функционируют на низких частотах и где окружающее пространство является ближним полем, взаимодействие осуществляется при помощи квазистатических волн. Системы RFID, функционирующие на высоких частотах, обычно работают в дальней зоне; при этом взаимодействие осуществляется при помощи электромагнитных волн.

Необходимо отметить, что в действительности граница между ближней и дальней зонами зависит от геометрии антенны. Мы будем использовать точку $r = \lambda / 2\pi$ в качестве стандартного определения границы между ближней и дальней зонами электромагнитного поля.

2.4. Импедансы пространства, антенн и цепей

Перед проведением дальнейшего анализа поведения систем RFID в ближней и дальней зонах поля нам будет полезно рассмотреть импедансы пространства, антенн и цепей. В общем физическое понятие импеданс описывает соотношения между воздействием и потоком (an effort and a flow).

В теории электромагнитных полей и волн импеданс описывает соотношение между электрическим и магнитным полями:

$$Z = E / H. \quad (2.22)$$

В антенне или электрической цепи это соотношение между напряжением и током:

$$Z = U / I. \quad (2.23)$$

Независимо от области использования импеданс Z исключительно полезный параметр, способный характеризовать поведение полей и волн, излучение и реакцию антенны, а также передачу мощности между антенной, фидером и нагрузкой.

На рис. 2.2а показаны импедансы свободного пространства, идеального электрического диполя и малой рамки. Мы видим, что после граничной точки ($\lambda/2\pi$) между ближней и дальней зоной, импедансы сближаются и становятся постоянными. Также видно, что в дальней зоне модули импедансов двух различных антенн совпадают и равны импедансу свободного пространства η . Однако в ближней зоне они различны.

На рис. 2.2б показана фаза импедансов двух антенн. Мы замечаем, что фазы в ближней зоне противоположны и равны $\pm\pi/2$, а в дальней зоне они приближаются к 0.

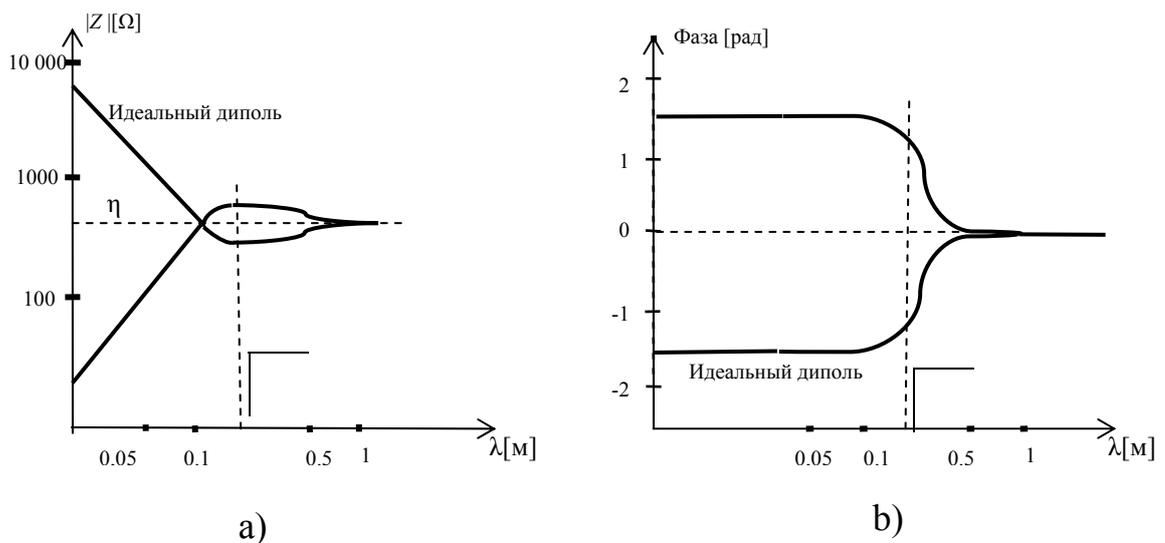


Рис. 2.2. а) – график модуля импедансов идеального электрического диполя и малой рамки; б) – график фазы импедансов идеального электрического диполя и малой рамки.

Антенны на своих входных зажимах имеют собственный импеданс. Реальная часть этого импеданса представляет собой комбинацию сопротивления излучения R_{rad} и омических потерь R_{ohmic} , а реактивная часть X – определяет энергию, запасенную в поле антенны:

$$Z_{\text{in}} = R_{\text{rad}} + R_{\text{ohmic}} + jX. \quad (2.24)$$

Антенны небольших по сравнению с длиной волны размеров имеют малую резистивную компоненту излучения и большую реактивную компоненту. Такое соотношение компонентов приводит к неэффективности излучения. Если мы вновь обратимся к рис. 2.2, то увидим, что малая рамка обладает большой положительной реактивной компонентой, а идеальный диполь – большой отрицательной реактивной компонентой. Положительная компонента соответствует индуктивности, а отрицательная – емкости. Это особенно важно для систем RFID, которые работают в ближней зоне.

Если размеры антенны увеличиваются по сравнению с длиной волны, резистивная компонента возрастает, а реактивная – уменьшается. При длине или периметре антенны, равной половине длины волны, реактивная компонента стремится к нулю, а резистивная - достигает максимума. При таких размерах антенна является резонатором и излучает эффективно.

Резонансные антенны обычно используются в системах, которые работают в дальней зоне. Дальнейшее увеличение размеров приводит к увеличению реактивной компоненты и уменьшению резистивной; при размерах, равных длине волны, импеданс близок к импедансу антенны с бесконечными размерами. Такой цикл повторяется с периодом в длину волны.

Фидерные линии и цепи тоже характеризуются импедансами. Из теории цепей известно, что для достижения максимума передачи мощности импедансы генератора и нагрузки должны быть комплексно сопряженными. При этом реальные компоненты этих импедансов должны быть равными, а реактивные – равными и противоположными. В следующих разделах мы увидим, как это влияет на работу систем RFID.

2.5. Связь в дальней и ближней зонах поля

Теперь мы рассмотрим антенны и основные принципиальные соотношения, необходимые для информационной и энергетической взаимосвязи считывателя и метки в дальней и ближней зонах поля. Благодаря отличиям поведения электромагнитного поля в каждой зоне, подходы существенно различны и мы рассмотрим их отдельно. В ближней зоне связь в системе RFID осуществляется между истоком и стоком энергии, тогда как в дальней зоне - между передатчиком и приемником.

2.5.1. Связь в ближней зоне поля

Как было показано ранее, электромагнитные поля в ближней зоне являются по своей сути реактивными и квазистатическими. Электрические поля не связаны с магнитными полями и, в зависимости от типа используемой антенны, одни могут преобладать над другими. В случае идеального диполя преобладают электрические поля, в то время как в случае малой рамки доминирует магнитное поле. Связь между меткой и считывателем может быть емкостной, при использовании электрического поля, или индуктивной – при использовании магнитного поля. Среди систем RFID ближнего поля индуктивно связанные системы находят более широкое распространение, чем емкостные. Поэтому мы более подробно будем рассматривать индуктивно связанные системы.

В этом разделе мы рассмотрим принципы и основные соотношения, касающиеся как индуктивных, так и емкостных систем.

2.5.1.1. Индуктивная связь

В антеннах индуктивных систем предполагается взаимодействие посредством квазистатического магнитного поля. Эти системы по существу представляют собой трансформатор, в котором ток, протекающий по его первичной обмотке, индуцирует магнитное поле, которое, в свою очередь, индуцирует ток и напряжение во вторичной обмотке (рис. 2.3). В случае RFID первичная обмотка размещена в считывателе, а вторичная – в метке. Небольшая рамочная антенна предпочтительнее дипольной антенны, так как в ближней зоне магнитное поле рамочной антенны значительно превышает магнитное поле дипольной антенны. Теперь важно пояснить различие между малой и большой рамками. Как указывалось ранее, периметр малой рамки меньше четверти длины волны. Ближнее поле и излучение в дальней зоне большой и малой рамок существенно различны. Распределение тока по периметру в большой рамке весьма неравномерно, тогда как в малой рамке оно приблизительно равномерно. Кроме того, в уравнениях поля малой рамки полагается, что расстояние от метки до рамки существенно больше ее размера, что может не всегда выполняться.

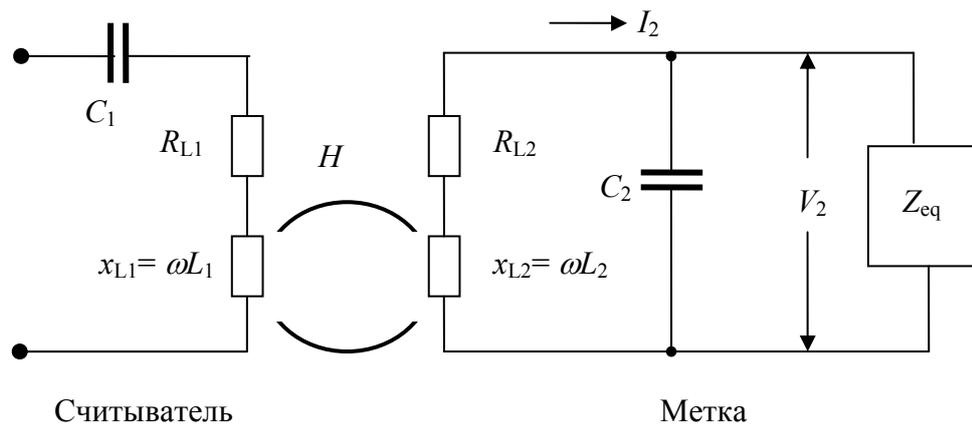


Рис. 2.3. Эквивалентная схема считывателя и рамки.

Прежде всего, для того чтобы определить мощность, индуцируемую во вторичной обмотке, необходимо рассмотреть уравнения магнитного поля в ближней зоне рамочной антенны. В случае малой рамочной антенны предположим, что расстояние от метки до рамки значительно больше ее радиуса.

С учетом того, что прямое вычисление уравнений для большой рамки является достаточно сложным, для подходящей аппроксимации магнитного поля используем закон Био-Савара. Закон Био-Савара непосредственно связывает магнитное поле H с распределением тока без вычисления векторного потенциала. Предполагая распределение тока в рамке из N витков равномерным, согласно закону Био-Савара получим:

$$\mathbf{H} = (NI/4\pi) \int_c \frac{d\mathbf{I} \times \mathbf{R}}{R^3}. \quad (2.25)$$

Когда расстояние до рамки порядка ее радиуса, при помощи последнего уравнения можно вычислить магнитное поле на оси, перпендикулярной плоскости рамки, где обеспечивается максимальная интенсивность поля, и, следовательно, максимальное расстояние считывания. В ближней зоне в случае малой рамки получим:

$$\mathbf{H} = 0,5N_I b^2 (z^2 + b^2)^{-3/2} \mathbf{z}, \quad (2.26)$$

где b – радиус рамки. Вычислив магнитное поле, мы можем определить его влияние на вторичную антенну. Для достижения максимальной мощности связи поля вторичная антенна должна перехватывать максимально возможное магнитное поле. В случае, когда магнитное поле постоянно,

взаимный поток между двумя антеннами через некоторую поверхность площадью A равен:

$$\Phi_{12} = \int_{S_2} \mathbf{B}_1 \cdot d\mathbf{s}_2 = \mu H_1 A_2 \cos \psi, \quad (2.27)$$

где ψ – угол между линиями поля и нормалью к плоскости. Взаимная индуктивность L_{12} может быть выражена через взаимный поток:

$$L_{12} = N_2 \Phi_{12} / I_1. \quad (2.28)$$

Согласно закону Ленца уменьшение этого взаимного потока во времени индуцирует напряжение во вторичной обмотке:

$$V_{1 \rightarrow 2} = -N_2 (d\Phi_{12} / dt) = j\omega N_2 \Phi_{12}. \quad (2.29)$$

Так как вторичная обмотка и подсоединенная к ней электрическая цепь имеют равные импедансы, напряжение $V_{1 \rightarrow 2}$ генерирует ток I_2 . Этот ток вызывает дополнительный магнитный поток, который противоположен начальному потоку, обусловленному собственной индуктивностью L_2 вторичной обмотки. Дополнительный поток вызывает падение напряжения $V_{2 \rightarrow 2}$:

$$V_{2 \rightarrow 2} = -L_2 (dI_2 / dt) = j\omega L_2 I_2. \quad (2.30)$$

Индуктивность малой круглой рамки равна [13]:

$$L_2 = \mu b N^2 [\ln(8b/a) - 1,75], \quad (2.31)$$

где b – радиус рамки, a – радиус провода рамки и $b \gg a$.

Вторичная обмотка (рамка) имеет сопротивление R_{L2} , которое вызывает дополнительное падение напряжения:

$$V_{RL2} = I_2 R_{L2}, \quad (2.32)$$

где омическое сопротивление малой круглой рамки [13]:

$$R_{L2} = N^2 (b/a) (\omega \mu / 2\sigma)^{1/2}. \quad (2.33)$$

Комбинируя напряжение взаимной индукции, напряжение самоиндукции и напряжение резистивной компоненты, получим:

$$V_2 = V_{1 \rightarrow 2} - V_{2 \rightarrow 2} - V_{RL2}. \quad (2.34)$$

Подставляя (2.27), (2.29), (2.30) и (2.32) в (2.34), получим:

$$V_2 = j\omega N_2 \mu_0 H_1 A_2 \cos \psi - I_2 (j\omega L_2 + R_{L2}) \quad (2.35)$$

Аналогичное соотношение справедливо для индукции из вторичной обмотки в первичную (для RFID – из метки в считыватель):

$$V_1 = j\omega N_1 \mu_0 H_2 A_1 \cos \psi - I_1 (j\omega L_1 + R_{L1}) \quad (2.36)$$

2.5.1.1.1. Резонанс и добротность

Чтобы максимизировать мощность во вторичной обмотке, необходимо как в метке, так и в считывателе создать резонансные цепи LC . Резонансная частота LC – цепи:

$$\omega_{LC} = (LC)^{-1/2}. \quad (2.37)$$

В считывателе для получения максимальной напряженности поля необходимо максимизировать выходной ток. Следовательно, чтобы импеданс на резонансной частоте стремился к нулю, последовательно с катушкой индуктивности следует включить емкость.

В метке для обеспечения питанием управляющей схемы необходимо максимизировать напряжение, поэтому, чтобы импеданс на резонансной частоте стремился к бесконечности, емкость следует включить параллельно катушке индуктивности. Так как обмотка, соединительные цепи и другие компоненты имеют резистивные составляющие, то считыватель и метка имеют конечную добротность Q . Добротность Q определяется как отношение запасенной энергии к энергии диссипативных потерь или как отношение центральной частоты к полосе пропускания по уровню 3 дБ. Она может быть выражена следующим образом:

$$Q = \omega L / R \quad \text{или} \quad Q = f_0 / \Delta f_{3\text{дБ}}. \quad (2.38)$$

Для того чтобы максимизировать Q , можно или уменьшить сопротивление, или увеличить индуктивность, либо сделать и то и другое. Для увеличения мощности связи, Q желательно делать максимальной. Однако увеличение Q соответствует узкой полосе пропускания, которая должна быть достаточно широкой для обеспечения необходимых

характеристик связи системы. Полоса пропускания как характеристика связи будет рассмотрена в главе 3. К этому вопросу мы также вернемся в главе 5.

2.5.1.1.2. Модуляция нагрузки

Мы увидели, что взаимосвязь между напряженностью магнитного поля и напряжением индукции используется как при передаче энергии и связи между считывателем и меткой, так и при связи между меткой и считывателем. Связь между меткой и считывателем обеспечивается за счет вариации импеданса метки. Изменяя этот импеданс, метка изменяет ток в первичной обмотке (антенне считывателя), что приводит к вариации магнитного поля. Эта вариация обнаруживается приемником считывателя.

Метка варьирует свой импеданс за счет изменения сопротивления или емкости; соответственно реализуется или омическая или емкостная модуляция нагрузки [6]. При омической модуляции резистор, параллельный нагрузке, включается или выключается либо одновременно с тактовой частотой данных, либо на другой, более высокой частоте (на поднесущей). Параллельное включение резисторов уменьшает общее сопротивление, что увеличивает ток метки. Это, в свою очередь, приводит к падению напряжения в считывателе. Таким образом обеспечивается амплитудная модуляция сигнала.

В случае емкостной модуляции, аналогично омической модуляции, включается или выключается конденсатор, параллельный нагрузке. Эти манипуляции обнаруживаются считывателем в виде амплитудной и фазовой модуляции сигнала.

2.5.1.1.3. Напряжение в нагрузке

В случае резонансного конденсатора, включенного параллельно эквивалентному импедансу подсоединенной схемы, выражение для напряжения в метке можно упростить. Обратившись к рис. 2.3, мы, прежде всего, определим ток I_2 в обмотке (антенне) метки в зависимости от импеданса параллельно включенных и настроенных в резонанс конденсатора и эквивалентного импеданса подсоединенной схемы:

$$I_2 = V_2 (Z_{C2}^{-1} + Z_{eq}^{-1})^{-1} = V_2 (Z_{eq}^{-1} + j\omega C_2). \quad (2.39)$$

Подставляя это выражение в (2.35) получим:

$$V_2 = j\omega N_2 \mu_0 H_1 A_2 \cos \psi [1 + (j\omega L_2 + R_{L2}) (Z_{eq}^{-1} + j\omega C_2)]^{-1}. \quad (2.40)$$

Это напряжение выделяется в нагрузке, которая подключена параллельно обмотке антенны. Ток в нагрузке можно найти простым применением закона Ома к импедансу нагрузки. Из выражения (2.40) видно, что напряжение в метке возрастает с числом витков обмотки, напряженностью магнитного поля, площадью петли и зависит от взаимной ориентации антенн. Все эти параметры определяют резонансную частоту и полосу пропускания метки. Эквивалентный импеданс нагрузки также влияет на резонансные свойства системы. Влияние параметров подсоединенной цепи не столь очевидно.

2.5.1.2. Емкостная связь

В системах с емкостной связью антенны считывателя и метки создают и взаимодействуют посредством квазистатического электрического поля. В таких системах напряженность поля определяется не токами, а распределением зарядов, что и определяет интенсивность взаимодействия. Так как интенсивность взаимодействия зависит не от тока, а от накопленного заряда, проводимость в этом случае играет менее важную роль, чем в индуктивных системах [15]. Емкостные системы, однако, более подвержены влиянию окружающих факторов, что и будет рассмотрено в последующих разделах.

Для систем с емкостной связью из-за преобладания электрического поля наиболее подходящей антенной является проводник или плоский диполь. В том случае, когда в качестве и первичной, и вторичной антенны используется диполь, может быть найдено грубое приближение электрического поля с использованием модели идеального диполя. Причем это может быть сделано только в предположении, что расстояние между первичной и вторичной антеннами много больше их размеров. На практике, однако, такой случай не реален. Поэтому мы определим электрическое поле проволочной дипольной антенны длиной L с равномерно распределенным зарядом ρ следующим выражением:

$$\mathbf{E} = (\rho / 4\pi \varepsilon) \int_L \frac{\mathbf{R}}{R^3} dL. \quad (2.41)$$

Используя закон Ампера, можно определить заряд на поверхности S :

$$Q = \int_S \mathbf{E} \cdot d\mathbf{S}. \quad (2.42)$$

Изменение заряда во времени индуцирует ток во вторичной антенне. Вторичная антенна имеет собственный импеданс, образованный емкостью электродов, активным сопротивлением антенны и других элементов цепи.

При этом индуцированный ток, соответствующий эквивалентному импедансу антенны и подсоединенной цепи, вызовет разность потенциалов на каждом элементе нагрузки.

2.5.1.2.1. Резонанс и модуляция нагрузки

В системах с емкостной связью, как и в системах с индуктивной связью, для обеспечения максимальной связи необходимы резонансные цепи. Так как антенна имеет собственную емкость, индуктивность включается параллельно в метке и последовательно в считывателе. Аналогично системам с индуктивной связью метка взаимодействует со считывателем при помощи вариации своего импеданса.

С целью обеспечения взаимодействия и передачи энергии метке от считывателя в емкостных системах, также как и индуктивных, желательно максимизировать связь. При использовании электромагнитных волн, цель та же, но механизмы обеспечения взаимодействия совершенно различны. В следующем разделе мы рассмотрим эти механизмы.

2.5.2. Связь в дальней зоне поля

В беспроводных коммуникационных системах, работающих в дальней зоне поля, взаимодействие между элементами этих систем достигается за счет передачи, распространения и приема электромагнитных волн. После рассмотрения некоторых полезных свойств электромагнитных волн остановимся на параметрах, которые важны для описания излучающих свойств антенны. Затем рассмотрим передачу и прием электромагнитных волн, обращая основное внимание на мощность, которая выделяется в антенне и подключенной к ней нагрузке.

2.5.2.1. Параметры антенн

Мы рассматривали элементарный электрический диполь и рамочную антенну. Эти антенны характеризуются высоким реактивным сопротивлением (либо индуктивным, либо емкостным), неэффективными излучающими свойствами и сложностью согласования. Будучи удобными для работы в ближней зоне, они практически не годятся для работы в дальней зоне, где передача и особенно прием энергии должны быть эффективными. По этой причине в дальней зоне чаще всего используются резонансные антенны, характерные размеры которых соизмеримы с длиной волны излучаемого сигнала. Резонансные антенны обладают более эффективными излучающими свойствами и низкой реактивной составляющей импеданса. Наиболее часто в системах RFID используются

полуволновые диполи и патчевые (микроросконовые) антенны. Некоторые физические характеристики таких антенн будут рассмотрены в главе 5.

Электромагнитные волны, излучаемые антеннами, представляют собой суперпозицию взаимосвязанных электрического и магнитного полей. Интенсивность электромагнитных волн зависит от типа антенны и выходной мощности генератора и убывает с увеличением расстояния от источника. Угловое распределение волн определяется только антенной.

Для описания углового распределения интенсивности поля мы будем использовать диаграмму направленности антенны. Диаграмма направленности антенны определяется нормированной электрической составляющей поля. В одномерном представлении диаграмма излучения описывается следующим образом:

$$F(\theta) = E_{\theta} / E_{\theta}(\max). \quad (2.43)$$

В случае, когда необходимо учитывать фазовые соотношения, за ноль фазы обычно принимается точка максимума электрического поля.

Антенны способны концентрировать поля в узкие пучки излучения. При этом увеличивается плотность потока мощности и дальность передачи. Для описания того как антенна концентрирует энергию в одном направлении по сравнению с любыми другими направлениями используется термин *направленность*.

Направленные свойства излучения определяется только диаграммой направленности антенны. Часто полезно, однако, описывать не только направленные свойства антенны, но и эффективность преобразования ее входной мощности в излучаемую выходную мощность. Для этого используется термин *усиление*. Иногда усиление определяется как отношение интенсивности излучения в данном направлении к мощности, которая подводится к антенне. Точнее коэффициентом усиления называется отношение мощности, подводимой к эталонной антенне, к мощности, подводимой к направленной антенне при условии получения одинаковой напряженности поля в точке приема:

$$G(\theta, \phi) = [E_A^2(\theta, \phi) / P_A] / (E^2 / P), \quad (2.44)$$

где E_A и E – напряженности полей, создаваемые направленной антенной и эталонной антенной, а P_A и P – мощности, подводимые к направленной антенне и эталонной антенне. Когда усиление обозначают просто G , имеют в виду его максимальное значение.

Итак, усиление определяют сравнением максимальной интенсивности излучения антенны с максимальной интенсивностью излучения некоторой эталонной (стандартной) антенны. Обычно измеряют усиление антенны по отношению к изотропному излучателю, который рассеивает энергию равномерно во всех направлениях. В отдельных случаях мы будем оценивать усиление по отношению к полуволновому диполю. Обычно усиление антенн измеряется в децибелах (dB). Изотропный излучатель имеет усиление 0 dB, а полуволновой вибратор – 2,15 dB. Когда рассматривают усиление относительно изотропного излучателя, единицу измерения записывают как dBi. Когда рассматривают усиление относительно полуволнового диполя, единицу измерения записывают как dBd. Когда мы рассматриваем антенну с усилением 6 dB, то это можно понимать как 6 dBi или как 3,85 dBd [13].

Таким образом:

$$G[\text{dB}] = G[\text{dBi}] = G_d[\text{dBd}] + 2,15 \text{ dB}. \quad (2.45)$$

В дальнейшем, когда мы будем использовать термин *усиление*, и если не сделано оговорок, будет иметься в виду усиление относительно изотропного излучателя.

Учитывая относительность определения усиления антенны, аналогичным образом описывают выходную мощность. Часто используют термин *эффективная (или эквивалентная) изотропная излучаемая мощность (effective (or equivalent) isotropically radiated power – EIRP)*, определяемый как мощность, подводимую к антенне, умноженную на ее усиление относительно изотропной антенны:

$$EIRP = G_t P_t. \quad (2.46)$$

Также используется термин *эффективная излучаемая мощность (effective radiated power – ERP)*, которая является мощностью на входе антенны, умноженной на ее усиление относительно полуволнового диполя:

$$ERP = G_{td} P_t. \quad (2.47)$$

Мощности EIRP и ERP связаны соотношением:

$$EIRP = ERP \times 1,64. \quad (2.48)$$

Используя эти параметры, мы рассмотрим передачу и прием электромагнитных волн.

2.5.2.2. Передача и прием сигнала

При рассмотрении работы меток RFID, мы должны определить мощность, принимаемую антенной метки. Часть этой мощности, необходимую для обеспечения собственных энергетических потребностей и обнаружения информации, метка поглощает. Другую часть мощности, необходимую для передачи информации считывателю, она отражает. Для того чтобы понять эти процессы, необходимо прежде всего определить мощность в метке, которая обеспечивается считывателем за счет излучения им энергии. Мы должны также определить мощность, поступающую в нагрузку метки. В итоге мы рассмотрим процесс отражение мощности обратно к считывателю, необходимый для осуществления связи метки и считывателя.

2.5.2.3. Передача сигнала от считывателя к метке

Падающая электромагнитная волна индуцирует на нагрузке антенны напряжение

$$V_A = \mathbf{E}^i \cdot \mathbf{h}^*, \quad (2.49)$$

где \mathbf{E}^i – напряженность падающей электромагнитной волны, \mathbf{h}^* – комплексно сопряженный вектор эффективной длины антенны (рис. 2.4). Обычно метки накапливают энергию и в дальнейшем используют ее по мере необходимости. Выходная мощность антенны метки должна быть достаточной для функционирования электронной схемы. Если мощность недостаточна, расстояние должно быть уменьшено.

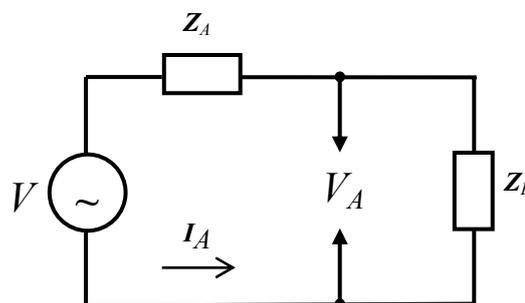


Рис. 2.4. Эквивалентная схема метки в дальней зоне.

Для передатчика с мощностью P_t и усилением антенны G_t падающая плотность потока мощности S на метке на расстоянии R равна:

$$S = P_t G_t / 4\pi R^2. \quad (2.50)$$

Эта плотность потока мощности определяется из вектора Пойнтинга, заданного уравнением (2.6), и вычисляется из электрического и магнитного полей. Из этой плотности потока мощности определяется эффективная площадь (апертура) антенны A_e , которая связана с собственным усилением антенны. Важно отметить, что эта площадь определяет мощность, которую антенна извлекает из электромагнитного поля:

$$A_e(\theta, \phi) = G(\theta, \phi) \lambda^2 / 4\pi. \quad (2.51)$$

При простом обозначении (без угловой зависимости) A_e обозначает максимальную эффективную площадь.

Перемножение эффективной апертуры на плотность потока мощности дает мощность, принимаемую приемной антенной. Необходимо, однако, учесть возможную поляризационную расстройку. Если антенна имеет поляризацию не совпадающую с поляризацией падающей волны, то принимается лишь часть мощности поля. Поэтому для того, чтобы учесть возможное поляризационное рассогласование, вводят коэффициент поляризационного рассогласования p [13]. Он определяется как:

$$p = |\mathbf{E}^i \cdot \mathbf{h}^*|^2 \cdot |\mathbf{E}^i|^2 \cdot |\mathbf{h}^*|^2, \quad (2.52)$$

где \mathbf{E}^i и \mathbf{h} определены в (2.49). При отсутствии поляризационного рассогласования коэффициент $p = 1$.

Теперь мы можем найти мощность, принимаемую антенной метки:

$$P_r = p S A_{er}(\theta, \phi) = p G_t P_t A_{er}(\theta, \phi) / (4\pi R)^2. \quad (2.53)$$

Выражая A_{er} через G_r с использованием уравнений (2.51) и (2.53) получим:

$$P_r = p P_t G_t G_r \lambda^2 / (4\pi R)^2. \quad (2.54)$$

Последнее уравнение показывает мощность, принимаемую антенной. Так как мы имеем подключенную к антенне нагрузку, мы должны определить мощность, выделяемую на нагрузке. Для того чтобы достичь максимума передачи мощности от антенны к нагрузке, необходимо

согласовать импеданс антенны с эквивалентным импедансом нагрузки. При комплексном сопряжении импедансов антенны и нагрузки мощность, выделяемая в нагрузке, равна мощности, принимаемой антенной. При рассогласовании нагрузки и антенны необходимо ввести коэффициент рассогласования импедансов q [13]:

$$q = P_D / P_{Dmax} = 4 R_A R_L / [(R_A + R_L)^2 + (X_A + X_L)^2]. \quad (2.55)$$

При комплексно сопряженных импедансах антенны Z_A и нагрузки Z_L коэффициент q равен 1. С использованием коэффициента q определим мощность, выделяемую в нагрузке R_L :

$$P_D = q P_r = q p P_t G_t G_r \lambda^2 / (4\pi R)^2. \quad (2.56)$$

Часть мощности, равная $(1 - q) P_t$, не выделяется в нагрузке, а отражается от нее и переизлучается антенной.

2.5.2.4. Модуляция отраженного сигнала

Большинство пассивных меток RFID, использующих энергию поля и в ближней, и в дальней зоне, достигают взаимодействия со считывателем за счет вариации импеданса нагрузки. При этом в дальней зоне вариация импеданса нагрузки достигается управляемой расстройкой импедансов нагрузки и антенны. Это приводит к тому, что часть мощности отражается обратно через антенну и рассеивается приблизительно так же, как антенна излучает свой собственный сигнал. Отраженный обратно сигнал обнаруживается и декодируется считывателем. Такая форма взаимодействия называется модуляцией отраженного сигнала (backscattering modulation). После краткого рассмотрения общих закономерностей отражения мы достаточно подробно рассмотрим отражение электромагнитных волн от антенн. Затем сравним мощность, которая принимается антенной и выделяется в нагрузке, с мощностью, отражаемой обратно к источнику излучения.

2.5.2.4.1. Отражение

Когда электромагнитная волна распространяется в среде и попадает на неоднородность, она рассеивается. Это явление называется отражением. В радиоэлектронных системах, которые содержат приемник, таких, например, как РЛС или RFID, передатчик излучает электромагнитную волну радиодиапазона, а приемник обнаруживает отражение от объекта. Когда приемник совмещен с передатчиком, отражение называется моностатическим или обратным отражением. Системы RFID достигают

взаимодействия с объектом за счет управляемого изменения обратного отражения метки.

Полезной характеристикой моностатического отражения объектов является поперечник обратного рассеяния или эффективная поверхность рассеяния (ЭПР). Сначала рассмотрим основные принципы модуляции обратного отражения применительно к системам RFID, а затем – более детально, факторы, влияющие на проектирование меток и считывателей этих систем.

2.5.2.4.2. Отражение и ЭПР

Организация IEEE определяет процесс отражения при распространении радиоволн как "процесс, при котором энергия распространяющейся волны рассеивается в результате взаимодействия с неоднородностями в среде" [16]. Это происходит в результате того, что электромагнитные волны падают на объект и индуцируют осциллирующие заряды и токи внутри и на поверхности объекта, а это, в свою очередь, приводит к появлению вторичной электромагнитной волны. Мы можем определить эти отраженные поля либо с использованием численного или аналитического вычисления индуцированных поверхностных зарядов и токов, либо с использованием аппроксимации тангенциальной составляющей поля, как поступают в физической оптике [17]. В общем, пространственное распределение отраженной энергии зависит от размеров, конфигурации и состава объекта, от длины волны и от направления ее прихода.

Эффективная поверхность рассеяния (ЭПР, англ. RCS - radar cross section) определяется как мера мощности, отраженной в определенном направлении. Она выражается через площадь, аналогично эффективной апертуре антенны. IEEE определяет ЭПР как:

«...умноженное на 4π отношение мощности в единице пространственного угла в определенном направлении к мощности на единице площади плоской волны, падающей на отражатель с определенного направления. Точнее, это предел этого отношения, когда расстояние от отражателя до точки измерения отраженной волны стремится к бесконечности ...» [16].

Рассматриваются три случая: моностатическая или ЭПР обратного отражения, когда направления падающей и отраженной волн совпадают, но противоположны по направлению, ЭПР прямого рассеяния, когда эти волны имеют одинаковое направление, а также бистатическая ЭПР, когда оба направления распространения волн различны. В дальнейшем, если не оговорено особо, используется моностатическая ЭПР. Также полагается, что

ЭПР является функцией частоты, поляризационных характеристик приемника и передатчика, а также угла визирования объекта:

$$\sigma = \lim_{R \rightarrow \infty} 4\pi R^2 (|\mathbf{E}^{\text{scat}}|^2 / |\mathbf{E}^{\text{inc}}|^2), \quad (2.57)$$

где \mathbf{E}^{scat} и \mathbf{E}^{inc} – отраженное и падающее электрическое поле соответственно, а R – расстояние от антенны приемника и передатчика до облучаемого объекта.

Возможно ЭПР определить как:

$$\sigma = 4\pi R^2 P_s / P_i, \quad (2.58)$$

где P_s и P_i – отраженная и падающая мощности соответственно.

При анализе отражений обычно рассматривают три области, в зависимости от соотношения длины волны λ к размерам отражающего тела L . Это следующие области отражений: Рэлеевская область (не следует относить это к плоской электромагнитной волне в дальней зоне, которая определяется выражением (2.21), резонансная область и оптическая область. В Рэлеевской области $\lambda \gg L$, так что фаза волны, падающей на тело, практически неизменна. При этом, поскольку тело воспринимает как бы квазистатическую волну, результирующее отраженное поле можно представить в виде индуцированного дипольного момента. В резонансной области размеры тела соизмеримы с длиной волны (обычно $L \cong 1 \dots 10 \lambda$). В этой области электромагнитная волна проявляет тенденцию «прилипания» к поверхности тела, создавая поверхностные волны: бегущие и краевые волны (traveling, creeping and edge traveling waves). В оптической области отражения длина волны значительно меньше размеров тела ($\lambda \ll L$). Здесь преобладающими механизмами отражения являются: зеркальное и краевое отражения, дифракция и многократные отражения [17].

Поскольку мы рассматриваем обратное отражение от антенн, размеры которых соизмеримы с длиной волны, будем полагать, что нашим случаем является резонансная область, где доминирует механизм отражения, связанный с поверхностными волнами. Мы не можем использовать плоскую тангенциальную аппроксимацию, так как она полезна в случаях зеркального отражения. В нашем случае для нахождения отраженной волны необходимо искать решение, основанное на использовании наведенных зарядов и токов. Аналитические решения чаще всего невозможны, используют численные методы, включая метод моментов [17].

2.5.2.4.3. Отражение от антенн

При анализе модулированных сигналов от меток в основном необходимо обращаться к отражающим характеристикам антенн меток. Так как антенны предназначены для приема и передачи излучения, обычно рассматривают две разновидности отражений.

Первый вид отражений, структурный, определяется тем, что антенна имеет определенную конфигурацию, размеры и материал, из которого она изготовлена. Второй вид – собственно антенный, определяется тем, что антенна предназначена для передачи радиочастотной энергии и имеет специфическую диаграмму направленности. Следует отметить, что отсутствует точное формальное определение этих видов отражений, и имеются их различные определения [13,17].

Существует несколько моделей, описывающих рассеянное поле антенны [18,19]. Мы будем использовать модель, представленную Грином [19]. Из рассмотрения рис. 2.4. видно, что отраженное поле зависит от импеданса нагрузки Z_L :

$$\mathbf{E}(Z_L) = \mathbf{E}(Z_A^*) + \Gamma I(Z_A^*) \mathbf{E}^r, \quad (2.59)$$

где Γ – модифицированный коэффициент отражения, равный

$$\Gamma = (Z_A^* - Z_L) / (Z_A + Z_L), \quad (2.60)$$

а \mathbf{E}^r - электрическое поле в дальней зоне, где антенну можно представить в виде элементарного источника тока:

$$\mathbf{E}^r = -j (\eta\beta / 4\pi r) e^{-j\beta r} \mathbf{h}. \quad (2.61)$$

Отрицательный знак указывает на наличие отражения. Сравнивая (2.61) с выражением для электрического поля идеального диполя в дальней зоне (2.13), видим, что в этом случае $\mathbf{h} = dl \sin\theta \boldsymbol{\theta}$ (для реальных антенн данное тождество несправедливо).

Члены $\mathbf{E}(Z_A^*)$ и $I(Z_A^*)$ в уравнении (2.59) представляют собой отраженное поле и ток антенны соответственно. Второй член в уравнении (2.59) представляет собой собственно отраженное поле антенны. Когда импеданс нагрузки согласован с антенным импедансом, этот член стремится к нулю и остается только первый член, который определяет структурную составляющую отражения антенны.

Аналитическое вычисление структурной составляющей (модой) сложно и, обычно, производится с использованием метода моментов. Если

пренебречь структурной модой (это возможно из-за того, что нас интересует не абсолютная величина отраженного поля, а его вариация), эффективная поверхность рассеяния антенны может быть выражена через известные антенные параметры:

$$\sigma_{ant} = p\Gamma^2 G^2(\theta, \phi) \lambda^2 / 4\pi, \quad (2.62)$$

где p – поляризационный коэффициент из уравнения (2.52), который учитывает поляризационное рассогласование между падающим полем и отражающей антенной. Кроме того, Грин рассматривает случай, когда нагрузка антенны принимает два значения с модифицированными коэффициентами отражения Γ_1 и Γ_2 . Это позволяет исключить структурную моду отражений, которая не зависит от нагрузки. Тогда модуляция отраженного поля становится равной

$$\Delta E = -j(\eta_0/4\lambda R_A) \mathbf{h}(\mathbf{h} \cdot \mathbf{E}^{inc}) (\Gamma_1 - \Gamma_2) e^{-j\beta r}/r. \quad (2.63)$$

Соответственно вариация ЭПР равна

$$\sigma_{\Delta} = p^2 |\Gamma_1 - \Gamma_2|^2 G^2(\theta, \phi) \lambda^2 / 4\pi. \quad (2.64)$$

Эти уравнения показывают основные закономерности модуляции обратного рассеяния. При изменении импеданса нагрузки, метка способна модулировать амплитуду или фазу отраженного ею поля.

2.5.2.4.4. Отраженная и принимаемая мощность

Поскольку вариация эффективного импеданса нагрузки приводит лишь к управлению антенной модой отражений, структурной модой отражений можно пренебречь. Однако для снижения шумов в приемном канале считывателя структурную моду желательно ограничивать. Реализовать такое ограничение трудно, так как это приводит либо к усложнению конструкции метки, либо к снижению эффективности приема.

Из уравнения (2.56) видно, что часть мощности поступает в нагрузку, а остальная часть мощности отражается. Мощность, поступающую в нагрузку, обозначим q , а отраженную мощность через $(1 - q)$. Значение Γ из уравнения (2.60) связано с q следующим образом:

$$1 - q = |\Gamma|^2. \quad (2.65)$$

Таким образом, становится очевидной взаимосвязь уравнений (2.62) и (2.56). Зная вариацию ЭПР, связанную с антенной модой, мы можем

определить отраженную мощность, поступающую на антенну считывателя. Исходя из определения ЭПР излучаемая мощность перехватывается эффективной площадью антенны метки, а затем изотропно излучается. Используя основную форму уравнения радиолокации

$$P_{\text{tr}} = p_t (P_t G_t^2 \lambda^2) \sigma / (4\pi)^3 R^4, \quad (2.66)$$

где p_t – коэффициент поляризационной расстройки между отраженной волной и антенной считывателя, получим отраженную мощность на входе приемно – излучающей антенны:

$$P_{\text{tr}} = p p_t \Gamma^2 (P_t G_t^2 G_r^2 \lambda^4) / (4\pi R)^4. \quad (2.67)$$

Таким образом, мы видим, как взаимосвязаны мощность, поступающая в нагрузку метки, и мощность, отраженная меткой и поступающая на вход считывателя.

2.6. Влияние окружающего пространства

До сих пор мы рассматривали поведение электромагнитных полей и волн в свободном пространстве и при этом предполагали, что атмосфера является однородной и не поглощающей, а объекты, окружающие и влияющие на процессы передачи и приема, отсутствуют. На практике, однако, окружающая среда не соответствует модели свободного пространства. Свойства канала распространения сигналов зависят от таких параметров, как влажность и температура среды, а также определяются взаимодействием с различными окружающими предметами. В этом разделе мы кратко рассмотрим эти аспекты. Здесь и в дальнейшем в главе 7 будут кратко рассмотрены интерференция и шумы, которые также относятся к влиянию окружающей среды.

2.6.1. Потери.

Вследствие того, что в ближней и дальней зонах поведение электромагнитного поля существенно различно, мы рассмотрим эти зоны отдельно.

2.6.1.1. Потери в ближней зоне

Поскольку в системах RFID, функционирующих в ближней зоне, напряженность поля убывает пропорционально $1/r^2$ или $1/r^3$, в отличие от убывания $1/r$ в дальней зоне, элементы влияния ближней зоны на сигнал относительно хорошо известны [20]. Имеется в виду, что на небольших

расстояниях можно пренебречь потерями интенсивности поля, искажениями диаграмм и влиянием пространства на другие общие характеристики системы. Индуктивные и емкостные системы ближней зоны, однако, сильно подвержены влиянию окружающих материалов и предметов – в особенности проводящих материалов.

Индуктивные системы оперируют магнитными полями, которые чувствительны к токам. Когда достаточно большие проводящие предметы подвергаются воздействию изменяющихся во времени магнитных полей, в них индуцируются токи. Эти токи, называемые вихревыми токами, противодействуют магнитному полю, которое индуцировало их, а также приводят к омическим потерям. Вихревые токи ослабляют магнитные поля и в случае систем RFID могут расстраивать антенны и уменьшать дальность действия. Особенно чувствительными являются системы с высокой добротностью. Чтобы предотвратить вихревые токи для экранировки больших проводников используются медно-порошковые материалы или ферриты [6]. В трансформаторах для снижения вихревых токов и омических потерь обычно применяют ферриты, имеющие низкую электропроводность.

Системы с емкостной связью также очень чувствительны к окружающим предметам. Так как емкостные системы осуществляют взаимосвязь при помощи электрического поля, их проблемы обусловлены искажениями разности потенциалов. Любые заземленные объекты притягивают линии электрического поля антенны и существенно изменяют конфигурацию зоны взаимодействия считывателя и метки.

2.6.1.2. Потери в дальней зоне и многолучевое распространение

В дальней зоне антенн излученные электромагнитные волны распространяются в окружающей среде. Напряженность поля убывает пропорционально $1/r$, что, собственно, обуславливает большую дальность взаимодействия. Следствием свободного распространения и слабого спада интенсивности является повышенная чувствительность систем RFID в дальней зоне к отражениям, рассеянию или дифракции собственного излучения или излучения других источников.

Потери в дальней зоне можно характеризовать крупномасштабными или мелкомасштабными неоднородностями. Крупномасштабные неоднородности проявляются в виде изменения напряженности поля на больших расстояниях. Мелкомасштабные неоднородности обычно проявляются в виде многолучевого распространения, когда волны, распространяющиеся от источника к приемнику различными путями, могут интерферировать и вызывать большие изменения интенсивности поля в точке приема. Мелкомасштабные неоднородности характеризуются быстрыми флуктуациями на небольших расстояниях. Потери,

обусловленные как мелкомасштабными, так и крупномасштабными неоднородностями, описываются соответствующими моделями [21].

Модели крупномасштабных потерь описывают затухание мощности сигнала с расстоянием от передатчика. Они модифицируют обычный закон обратного квадрата в уравнении для свободного пространства (2.35) и описывают затухание, обусловленное атмосферой и взаимодействием с материалами. Полагая затухание равным

$$PL(R) = (\lambda / 4\pi R)^n, \quad (2.69)$$

где n равно 2 для свободного пространства, уравнение (2.41) преобразуется к виду:

$$P_r = p P_t G_t G_r PL(R). \quad (2.70)$$

Обычно используемая модель для оценки потерь распространения в пространстве внутри помещений задается логарифмической моделью [21]:

$$PL(R) [\text{дБ}] = PL(R_\sigma) + 10 n \log (R/R_0) + X_\sigma \quad (2.71)$$

где n зависит от свойств помещений, а X_σ – нормально распределенная переменная со стандартным отклонением σ , определенным в децибелах. Величина R_0 является фиксированным расстоянием, на котором проводятся измерения, и обычно выбирается равным 1 метру в условиях распространения поля внутри помещений. Таблица 2.1 показывает значения n и σ в различных условиях распространения поля и на различных частотах [23]. Меньшие значения σ соответствуют более точной модели.

Таблица 2.1. Параметры n и σ логарифмической модели потерь распространения в пространстве внутри помещений [23].

Здания	Частота (МГц)	n	σ
Розничный магазин	914	2,2	8,7
Бакалейный магазин	914	1,8	5,2
Плотный офис	1500	3,0	7,0
Неплотный офис	900	2,4	9,6
Неплотный офис	1900	2,6	14,1
Текстильные, химические предприятия	1300	2,0	3,0
Текстильные, химические предприятия	4000	2,1	7,0 – 9,7

Книжный магазин, булочная	1300	1,8	6,0
Металлообрабатывающие предприятия	1300	1,6 – 3,3	5,8 – 6,8
Стены жилых помещений	900	3,0	7,0

Явления многолучевого распространения описывают маломасштабные фединговые модели. Многолучевое распространение на небольших расстояниях может вызывать сильные флуктуации и амплитуды и фазы, случайные частотные модуляции и временную дисперсию, обусловленную задержками. В пространстве, содержащем металлические предметы и отражающие объекты, может иметь место многолучевое распространение и интерференция сигналов.

Для описания многолучевого распространения используются различные статистические модели, а общей мерой является среднеквадратическое отклонение (СКО) задержки распространения. Здания с небольшим числом металлических фрагментов и плотной планировкой обычно имеют небольшое СКО задержки распространения в пределах от 30 до 60 нс. Большие здания с большим числом металлических фрагментов и открытыми боковыми пристройками могут иметь задержки распространения, достигающие 400 нс. [22]. Технологиями, которые позволяют минимизировать эффект многолучевого распространения, являются эквалайзирование, многоканальность и канальное кодирование. В случае RFID в считывателях используется эквалайзирование и антенная многоканальность. В метках из-за жестких ограничений размеров, сложности и стоимости такие способы не используются.

2.6.2. Влияние соседних антенн

Когда множество меток находится на близком расстоянии друг от друга, связь между их антеннами может оказывать вредное влияние на прием и передачу сигналов.

В ближней зоне близко расположенные метки могут вызывать расстройку соседних антенн. Особенно чувствительными являются метки с высокой добротностью. Один из способов решения этой проблемы в системах ближнего поля состоит в том, что антенны меток, которые предполагается эксплуатировать близко друг к другу, следует настраивать на более высокую частоту.

В системах RFID дальней зоны диаграммы направленности антенн меток также могут существенно искажаться при их близком взаимном расположении и, соответственно, эффективность передачи мощности и сигналов будет снижаться.

2.6.3. Температура и влажность

Изменения температуры окружающей среды вызывают изменения параметров согласующих цепей и, следовательно, могут приводить к неэффективной передаче мощности. При этом системы с высокой добротностью из-за сдвига резонансной частоты могут подвергаться серьезной расстройке. Поэтому максимально по возможности должны применяться компоненты с низким температурным коэффициентом.

Влажность также может приводить к деградации характеристик. В общем случае эти эффекты наиболее вредны на высоких частотах.

2.7. Резюме

В этой главе мы рассмотрели, как при помощи электромагнитных полей и волн системы RFID достигают энергетической и информационной связи. Мы рассмотрели, как при помощи антенн создаются электромагнитные поля и волны, а также изучили свойства ближней и дальней зоны поля, излученного антенной. Мы поняли значение энергетической связи и информационной связи в каждой из этих зон. В заключение мы рассмотрели некоторые аспекты влияния окружающей среды на поведение полей и волн.

В следующей главе мы увидим, как посредством изменений параметров электромагнитных волн может передаваться информация и каким образом это ограничивает проектирование и функционирование систем RFID.

Глава 3

СВЯЗЬ

3.1. Введение

Связь между считывателем и меткой является неотъемлемой частью технологии RFID. Ранее нами в главе 2 были рассмотрены основы электродинамики, при помощи которых реализуется связь. Теперь мы рассмотрим собственно принципы связи.

После краткого рассмотрения методов анализа систем связи мы рассмотрим способы кодирования и модуляции, используемые в процессе связи в системе RFID. Завершим эту главу рассмотрением проблем надежности передачи данных, включая вероятность битовых ошибок (BER) и методы обнаружения ошибок.

При передаче информации нас, как правило, будут интересовать три основных параметра: полоса пропускания, вероятность ошибки, а также сложность и издержки обнаружения ошибок [23]. На протяжении всей этой главы мы будем акцентировать внимание на этих трех параметрах.

3.2. Процесс связи

Процесс связи состоит из передачи и приема информации. Для передачи по каналу связи с помехами информация преобразовывается, накладывается на несущий сигнал и передается. После приема сигнала с помехами, смесь демодулируется и обрабатывается для выделения первоначальной информации.

В технологии RFID информация включает команды управления и данные в двоичном виде. Обычно команды выделяются отдельными строками двоичных данных, но в некоторых случаях они представляют собой некоторую уникальную сигнатуру модуляции. В таких случаях необходима определенная обработка сигнала. Что касается данных, для них оказывается удобным использовать кодирующие устройства.

После того как данные закодированы, они налагаются на несущий сигнал. Такой процесс называется модуляцией. Модуляция необходима как для передачи данных при помощи канала распространения сигнала, так и для согласования спектра частот с административными регламентными ограничениями.

3.3. Сигналы и спектры

В этом разделе мы рассмотрим основы анализа сигналов. Мы обсудим такие понятия, как усреднение сигнала по времени, его мощность и

среднеквадратическое значение. Затем мы рассмотрим спектры и спектральную плотность мощности различных видов сигналов. Это будет полезно для понимания процесса кодирования и методов модуляции, рассматриваемых в последующих разделах.

3.3.1. Усреднение по времени, мощность и среднеквадратическое значение (СКО)

Перед рассмотрением спектральных характеристик сигналов, нужно понять различные методы их описания во времени. Мы кратко рассмотрим понятия усреднения по времени, приведенной (нормализованной) мощности и среднеквадратического значения сигнала.

Усредненное по времени или среднее значение (постоянная составляющая) сигнала $\omega(t)$:

$$\langle \omega(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \omega(t) dt. \quad (3.1)$$

Для периодического сигнала формула (3.1) упрощается:

$$\langle \omega(t) \rangle = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \omega(t) dt, \quad (3.2)$$

где T_0 – период сигнала.

Независимо от того, в каких единицах измерения описан сигнал $w(t)$ - вольт, ампер, вольт/м или ампер/м, его возведение в квадрат дает приведенную мощность в ваттах (соответствующая резистивная компонента принимается равной единице). Приведенная средняя мощность сигнала равна

$$P = \langle \omega^2(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \omega^2(t) dt. \quad (3.3)$$

Квадратный корень из приведенной средней мощности дает среднеквадратичное значение (СКВ) сигнала:

$$\sigma = \sqrt{\langle \omega^2(t) \rangle}. \quad (3.4)$$

Понятия временного усреднения, приведенной мощности и СКВ будут важны при рассмотрении спектров сигналов.

3.3.2. Преобразование Фурье и спектры

Каждый сигнал может быть представлен в виде спектра частот. Эти частоты позволяет определить преобразование Фурье. Такое определение полезно не только для согласования с регламентами, но также и для согласования частотных характеристик компонентов системы. Антенны, например, имеют определенные частотные (спектральные) характеристики, в соответствии с которыми они могут эффективно излучать или принимать сигналы. Передача или прием компонентов частоты сигнала, находящихся вне этой спектральной характеристики, будут невозможны. Зачастую антенны используются в качестве элементов фильтрации. В этом разделе мы кратко рассмотрим принципы и термины, которые используются для описания сигналов и спектров частот, связанных с ними.

3.3.2.1 Прямоугольный импульс

Если мы рассматриваем прямоугольный импульс во временной области $w(t)$ с длительностью импульса T и амплитудой A (рис. 3.1.а), то для его представления в частотной области используют преобразование Фурье [24]:

$$W(f) = \int_0^T A e^{-j2\pi ft} dt = \frac{jA}{2\pi f} (e^{-j2\pi fT} - 1) = A T e^{-j\pi fT} \left(\frac{\sin \pi fT}{\pi fT} \right). \quad (3.5)$$

Эта функция имеет размерность величины, определенной на временной оси, деленной на единицу частоты, то есть на герц. В случае если $u(t)$ – напряжение, то $W(f)$ имеет размерность вольт/герц. Величина (модуль) спектра прямоугольного импульса в частотной области имеет вид

$$|W(f)| = \frac{A}{2\pi|f|} [2 - 2\cos(2\pi fT)]^{\frac{1}{2}}. \quad (3.6)$$

Модуль спектра прямоугольного импульса представлен на рисунке 3.1b.

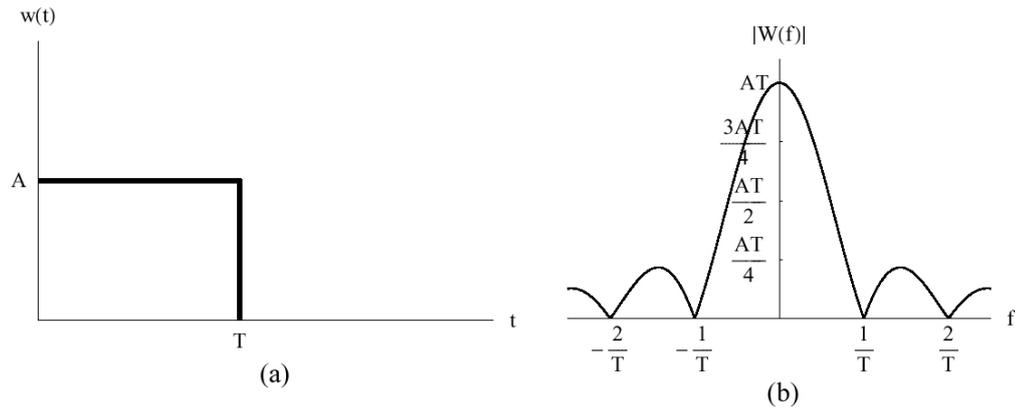


Рис. 3.1. (a) – эюра прямоугольного импульса во времени, (b) – спектр прямоугольного импульса.

На частотной оси при каждом кратном числе, которое обратно длительности импульса, имеются нули, а также пики, находящиеся на равном расстоянии между нулями. Пик в нуле имеет величину AT_r , другие пики имеют величины:

$$2AT_r/n\pi, \quad |n| = 3, 5, 7, \dots \quad (3.7)$$

Общая площадь под кривой огибающей спектра дает среднее значение сигнала в единицах его измерения.

Зная спектр сигнала, мы можем найти его спектральную плотность мощности S (СПМ, англ. Power Spectral Density – PSD) за период T :

$$S = \lim_{T \rightarrow \infty} \left(\frac{|W_T(f)|^2}{T} \right). \quad (3.8)$$

Спектральная плотность мощности (СПМ) прямоугольного импульса показана на рис. 3.2.

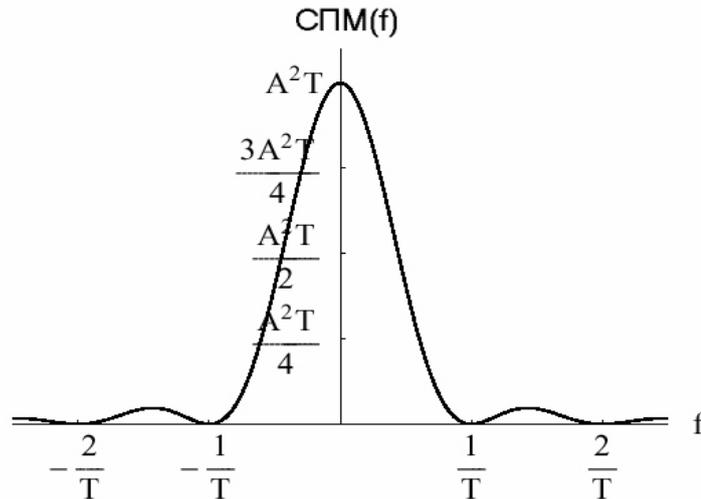


Рис. 3.2. Спектральная плотность мощности прямоугольного импульса.

Общая площадь под кривой спектральной плотности мощности соответствует усредненной по времени мощности сигнала.

3.3.2.2. Последовательность прямоугольных импульсов

Если мы рассматриваем в течение всего времени последовательность прямоугольных импульсов длительностью T_r и периодом T_0 (рис. 3.3), то его спектр можно определить при помощи ряда Фурье.

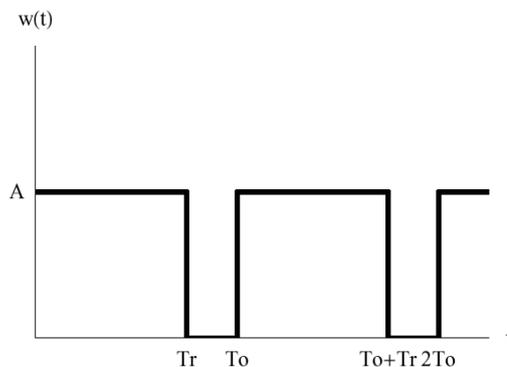


Рис. 3.3. Последовательность прямоугольных импульсов длительностью $T_r = 3/4 T_0$.

Спектр периодического сигнала равен:

$$W(f) = \sum_{n=-\infty}^{n=\infty} c_n \delta(f - nf_0), \quad (3.9)$$

где $f_0 = 1/T_0$. Свяжем T_0 с T_r при помощи постоянной r , которая определяется из $T_r = r \cdot T_0$. Тогда коэффициенты Фурье c_n в комплексном виде определяются следующим образом:

$$c_n = \frac{1}{T_0} \int_0^{rT_0} A e^{-j2\pi f_0 n t} dt = j \frac{A}{2\pi n} (e^{-j2\pi n r} - 1). \quad (3.10)$$

Формулу (3.10) можно записать в ином виде:

$$c_n = \begin{cases} Ar, & n = 0, \\ j \frac{A}{2\pi n} (e^{-j2\pi n r} - 1), & n \neq 0. \end{cases} \quad (3.11)$$

Модуль коэффициентов Фурье:

$$|c_n| = \begin{cases} Ar, & n = 0, \\ \frac{A}{2\pi |n|} [2 - 2 \cos(2\pi n r)]^{\frac{1}{2}}, & n \neq 0. \end{cases} \quad (3.12)$$

Модули спектров при $r = 0,25$, $r = 0,5$ и $r = 0,75$ представлены на рис. 3.4. Спектральные линии появляются на частотах, кратных f_0 , а их величина изменяется по кривой, соответствующей огибающей непрерывного спектра одиночного прямоугольного импульса длительностью T_r .

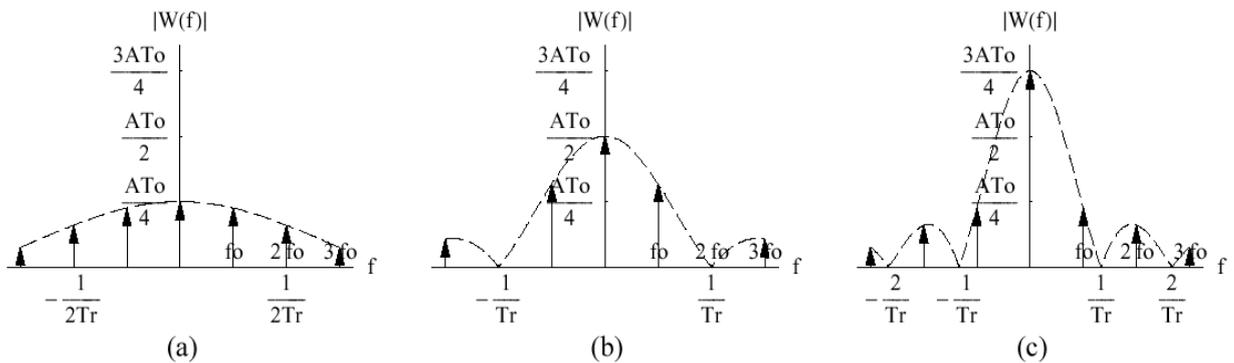


Рис. 3.4. Модули спектра для трех отношений T_r к T_0 : (a) $1/4$, (b) $1/2$ и (c) $3/4$.

Ширина полосы частот, которую занимает спектр, изменяется соответственно огибающей кривой непрерывного спектра. Видно, что с уменьшением длительности импульса, величина спектральных линий падает, но все они распределены в более широкой полосе частот. Это может привести к сложностям с приемом таких сигналов узкополосными антеннами, при этом может уменьшиться соотношение сигнал/шум. Также

возможны проблемы, связанные с превышением допустимой регламентами ширины спектра излучения.

Спектральная плотность мощности периодических импульсов может быть найдена при помощи возведения в квадрат модуля коэффициентов Фурье:

$$S(f) = \sum_{n=-\infty}^{n=\infty} |c_n|^2 \delta(f - nf_0). \quad (3.13)$$

Сумма спектральных линий от $n = -\infty$ до $n = +\infty$ соответствует средней мощности сигнала.

3.3.2.3 Фильтрация

Другим очень важным принципом оценки систем связи является принцип линейной инвариантности [24]. Линейно-инвариантные системы (ЛИС) – это системы, которые являются линейными, то есть подчиняющимися принципу суперпозиции, и инвариантными во времени, в которых временная задержка входного сигнала соответствует такой же временной задержке выходного сигнала. При помощи этих свойств мы можем оценивать более сложные сигналы, включая эффекты фильтрации, а также современные методы модуляции. Коротко опишем применение принципов ЛИС для фильтрации.

Когда на вход этой линейной системы, например, фильтра, подается импульс (дельта функция), на ее выходе будет сигнал, соответствующий импульсной характеристике этой системы. Импульсная характеристика системы удобна, поскольку она позволяет непосредственно определять выходной сигнал данной системы по входному. Свертка входного сигнала $x(t)$ с реакцией на импульсное возмущение $h(t)$ дает выходной сигнал $y(t)$. В частотной области операция свертки соответствует умножению, поэтому:

$$Y(f) = X(f) H(f), \quad (3.14)$$

где $H(f)$ называется частотной характеристикой или передаточной функцией системы, поскольку она преобразует входной спектр $X(f)$ в выходной спектр $Y(f)$.

Теперь, если мы рассмотрим последовательность прямоугольных импульсов (3.9) и пропустим ее через некоторую систему с частотной характеристикой $H(f)$, мы найдем:

$$Y(f) = \sum_{n=-\infty}^{n=\infty} c_n H(nf_0) \delta(f - nf_0). \quad (3.15)$$

В случае RC фильтра низких частот мы с некоторым упрощением можем определить его частотную характеристику при помощи законов Кирхгофа и преобразования Фурье. Частотная характеристика RC фильтра:

$$H(f) = \frac{1}{1 + j(2\pi RC)f}. \quad (3.16)$$

Подстановка в формулу (3.14) позволяет получить:

$$Y(f) = \sum_{n=-\infty}^{n=\infty} c_n \frac{1}{1 + j(2\pi RC)nf_0} \delta(f - nf_0), \quad (3.17)$$

где коэффициенты c_n определяются формулой (3.10). Рассматривая последовательность прямоугольных импульсов с $r = 0,5$ и постоянной RC , равной $T_r/6$, мы можем представить сигнал $u(t)$ и модуль спектра на выходе фильтра $|Y(f)|$ как показано на рис.3.5. Мы видим, что высокочастотные компоненты спектра опустились ниже аналогичных компонент спектра прямоугольных импульсов до фильтрации, приведенных на рис. 3.4.

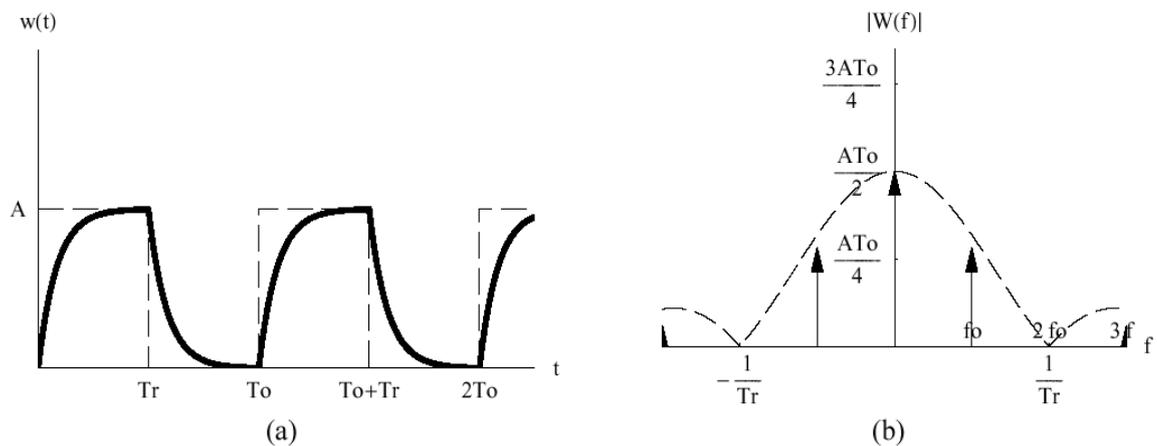


Рис. 3.5. Последовательность прямоугольных импульсов на выходе фильтра низких частот: (a) – сигнал, (b) – модуль спектра.

3.3.2.4. Модуляция

Сигналы, которые мы рассмотрели к настоящему времени, были видеосигналами, спектры которых сосредоточены около нулевой частоты. При передаче информации по беспроводным каналам необходимо при помощи модуляции перенести видеосигнал на частоту несущего сигнала. Тем самым создается модулированный сигнал со спектром, сосредоточенным около несущей частоты.

Перед рассмотрением модуляции в последующих разделах полезно исследовать общие спектральные характеристики модулированного сигнала. Полосовой сигнал с несущей частотой ω_c , модулированный видеосигналом $g(t)$ можно представить в виде [24]:

$$v(t) = \operatorname{Re}\{g(t)e^{j\omega_c t}\} = g(t)\cos(\omega_c t). \quad (3.18)$$

Спектр такого сигнала:

$$V(f) = \frac{1}{2}[G(f-f_c) + G(-f-f_c)], \quad (3.19)$$

а его спектральная плотность мощности:

$$S_v(f) = \frac{1}{4}[S_g(f-f_c) + S_g(-f-f_c)]. \quad (3.20)$$

Как показывает выражение (3.19), процесс модуляции сосредотачивает спектр полосового сигнала вокруг положительной и отрицательной несущей частоты, при этом величина спектральной линии каждой составляющей делится пополам. Аналогичный результат виден из выражения (3.20) для спектральной плотности мощности сигнала.

3.3.2.5. Ширина спектра

Представление о ширине спектра является особенно важным с точки зрения скорости передачи данных, соотношения мощности сигнала к мощности шума и удовлетворения нормам регламентов. Вообще, ширина спектра – это область положительных частот, которые занимает сигнал, однако однозначного определения не существует. Мы дадим те некоторые определения, с которыми столкнемся в дальнейшем [24]:

- Абсолютная ширина спектра – диапазон частот, вне которого уровень спектральных составляющих равен нулю.
- Ширина спектра по уровню n дБ, где n – некоторое значение, обычно равное 3, 20, или 60 дБ – это диапазон частот, где уровень спектральных составляющих не опускается ниже максимального значения, равного n дБ.
- Ширина спектра по первым нулям для систем ограниченной полосой пропускания – диапазон частот между первыми нулями выше и ниже центральной частоты.

При рассмотрении параметров сигнала в относительных единицах можно не указывать определение ширины спектра. Однако при рассмотрении сигналов в абсолютных единицах должно быть дано конкретное определение ширины спектра.

3.4. Кодирование в линии передачи сигнала

В цифровых системах связи символы двоичных данных должны быть преобразованы в форму, которая была бы удобной при их передаче. Обычно это делается при помощи передачи последовательности импульсов, которые форматированы для представления символов данных. Такое импульсное форматирование часто называется линейным кодированием. Кратко рассмотрим линейное кодирование с точки зрения беспроводной передачи данных в технологии RFID.

Линейное кодирование возникло в телефонии, где передача осуществлялась по линиям из медных проводов [25]. Однако такое кодирование находит широкое применение и в беспроводной передаче данных. На рис.3.6. показаны обычно используемые методы кодирования. В зависимости от исполнения они могут быть двух или трех уровневые – положительный и нулевой уровень, или положительный, отрицательный и нулевой уровни. Те сигналы, которые используют положительный и отрицательный уровни, называют биполярными или полярными, а сигналы, которые используют только положительный и нулевой уровни, относятся к униполярным. Поскольку линейное кодирование используется в проводной, узкополосной передаче данных, выбор уровней имеет существенное значение.

Так как в системах RFID передаются узкополосные сигналы, закодированная последовательность импульсов модулирует несущую частоту. В зависимости от используемой схемы модуляции может быть необходимым преобразование биполярного к униполярному сигналу, как, например, в случае амплитудной манипуляции. Мы не будем подробно анализировать кодирование до тех пор, пока не рассмотрим процесс модуляции, поскольку и кодирование и модуляция совместно определяют полосу пропускания, вероятность ошибки и сложность приемника. Тем не менее, мы рассмотрим спектральные плотности мощности интересующих нас способов кодирования.

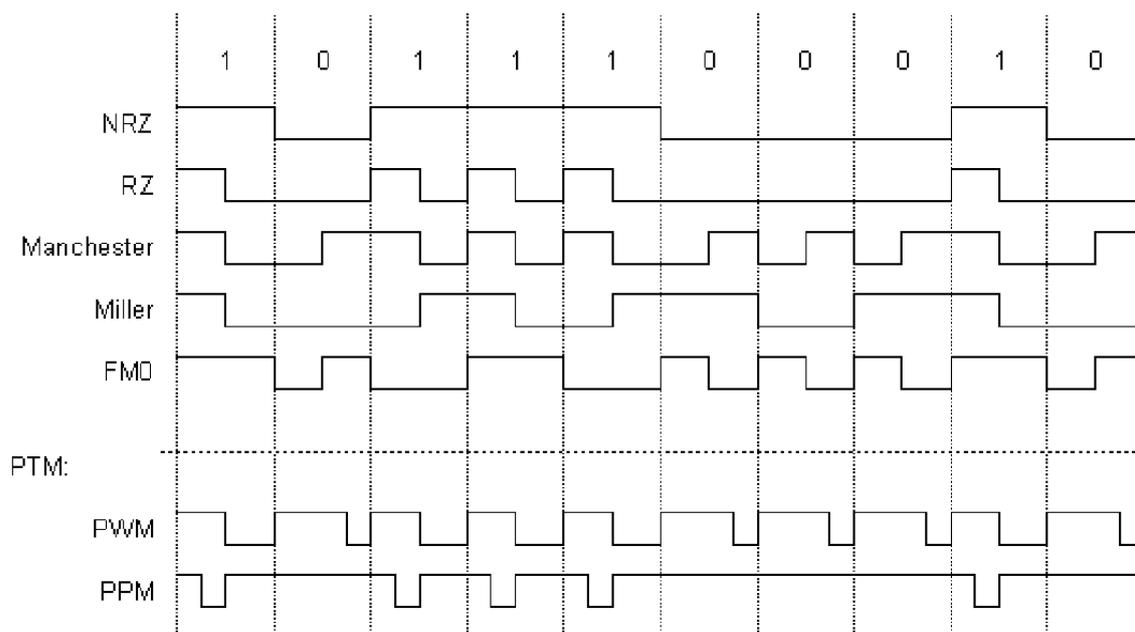


Рис. 3.6. Способы кодирования, используемые в системах RFID.

Существует два основных класса линейных кодов: уровневые коды и переходные коды [25]. Уровневые коды содержат информацию в уровне сигнала и данные обычно декодируются мгновенно, независимо от предшествующих данных. Примерами подобных уровневых кодов являются код без возврата к нулю (Non Return to Zero – NRZ) и с возвратом к нулю (Return to Zero – RZ). Переходные коды с другой стороны, содержат информацию в переходах сигнала. Сами переходы могут быть мгновенными, но они определяются на основе данных памяти, которая фиксирует закодированные данные. Обычно в технологии RFID используются такие коды, как Манчестерский код (код с переходом фазы) – Manchester, код Миллера (код с модуляцией задержки) – Miller, и код FM0 (двухфазово-пространственный код) – это все линейные коды. Коды Миллера и FM0 используют память о предыдущем состоянии кода, а Манчестерский код – не использует. Время-импульсная модуляция (ВИМ, англ. PTM – Pulse Time Modulation) для передачи информации варьирует длительность импульса и подразделяется на широтно-импульсную модуляцию (ШИМ, англ. PWM – Pulse Width Modulation) и фазо-импульсную модуляцию (ФИМ, англ. PPM – Pulse Position Modulation). Это не линейные коды и такие способы кодирования также используются при кодировании узкополосных сигналов.

Способ Манчестерского кодирования, известный также как расщепленно-фазовое кодирование или цифровой двухфазный код, является мгновенным переходным кодом. Название двухфазный происходит из-за того, что по существу прямоугольная волна со сдвигом фазы 0 градусов

соответствует 1, а прямоугольная волна со сдвигом фазы 180 градусов соответствует 0. Такая ситуация может вызывать сложности при приеме и, кроме того, этот код по сравнению с другими кодами имеет более широкий спектр и не способен обнаруживать ошибки. Тем не менее Манчестерский код широко используется в локальных сетях Ethernet. Он удобен потому, что, несмотря на биполярную форму, не имеет постоянной составляющей.

Кодирование Миллера – другой переходный код, который, в отличие от Манчестерского кода, требует наличия памяти о предыдущей информации. Он обеспечивает хорошее временное распределение, причем, по сравнению с Манчестерским кодом, без расширения спектра. Кроме того код Миллера более удобен для модуляции, так как в этом случае отсутствуют сложности с переходом фазы.

Кодирование FM0, также известное как двухфазово - пространственный код, является еще одним переходным кодом, который требует наличия памяти информации. Кодирование FM0 похоже на кодирование Миллера, однако в этом случае переход происходит точно в начале каждого битового периода. Это приводит к более легкому обеспечению синхронизации при приеме, однако, за счет расширения спектра.

Время-импульсная модуляция (ВИМ) – это такой класс способов кодирования, который обычно используется для кодирования аналогового сигнала в цифровой на временной шкале [24]. ВИМ включает широтно-импульсную модуляцию (ШИМ) и фазо-импульсную модуляцию (ФИМ). Такая модуляция часто используется в системах RFID для кодирования цифровых сигналов при временной передаче данных. В кодировании ФИМ наличие модуляции на некотором интервале (фрейме) соответствует 1, а отсутствие модуляции соответствует 0. В широтно-импульсной – ШИМ модуляции, двоичные единицы и нули могут быть представлены различными длительностями импульса за определенный интервал времени. Такие способы кодирования данных могут быть выгодными с точки зрения обеспечения заданной полосы пропускания, улучшения синхронизации и снижения вероятности ошибок.

Рассмотренные коды, обычно используемые в системах RFID – это только некоторая часть кодов, большинство из которых были разработаны для проводных систем. Так как в проводных и беспроводных системах процессы передачи и приема данных, а также каналы распространения сигналов существенно различны, то для конкретной области применения необходимо определить конкретный способ кодирования. При этом коды ВИМ обеспечивают необходимую гибкость выбора.

В следующем разделе мы проанализируем различные способы кодирования сигналов в сочетании с модуляцией.

3.4.1. Спектральная плотность мощности сигналов

Хотя переносчиком информации в беспроводной связи является модулированный несущий сигнал, для исследования его спектральных характеристик значительно более удобен собственно модулирующий кодированный сигнал. В зависимости от используемого типа модуляции, спектральные характеристики, вероятно, останутся теми же самыми, а спектр будет сосредоточен около несущей частоты и уменьшен по величине согласно формуле (3.19). Как было упомянуто выше, многие из видов кодирования первоначально были разработаны в биполярном варианте. Тем не менее, большая часть исследований их спектральных характеристик основана на использовании униполярной формы. По причинам, отмеченным в разделе 3.4, мы также будем вычислять спектральную плотность мощности кодированных сигналов в униполярной форме.

Есть два подхода к вычислению СПМ кодированных сигналов – детерминированные методы и вероятностные методы. Детерминированные методы вообще и, в частности, те, которые представлены в разделе 3.2.2, требуют обязательного знания точной формы сигнала. В вероятностных методах назначаются вероятности элементарных сигналов, из которых состоит сигнал. Поскольку в системах RFID осуществляется прием случайных идентификационных сигналов, мы будем использовать вероятностные методы анализа СПМ.

Часто случайный процесс, а в нашем случае это сигнал на выходе приемника, имеет некоторую специфическую структуру, которая может характеризоваться относительно небольшим количеством ограничений. Процесс Маркова – один из таких процессов. Процессы Маркова описывают случайные процессы, где будущее состояние зависит только от настоящего состояния, а не от прошлой истории. В общем, такая ситуация присуща всем кодированным сигналам, которые мы будем рассматривать. Когда имеется комбинация дискретных состояний, процесс Маркова называется цепью Маркова [26].

Следуя приближению [27], которое основано на [28], мы рассмотрим Марковский сигнал, который состоит из N элементарных сигналов (элементов) с набором $\{s_i(t); i = 1, 2, \dots, N\}$ и общим периодом T_s . Последовательность сигналов характеризуется набором стационарных вероятностей $\{p_i; i = 1, 2, \dots, N\}$ и набором вероятностей перехода $\{p_{ik}; i = 1, 2, \dots, N\}$. Стационарные вероятности показывают вероятность появления определенного элемента, в то время как вероятности перехода показывают вероятность появления элемента s_k после элемента s_i . Формально вероятности перехода определяются следующим образом:

$$p_{ik} = P\{s_{n+1} = k | s_n = i\}, \quad i, k \geq 1. \quad (3.21)$$

Эти вероятности могут быть выражены в виде матрицы вероятностей \mathbf{P} :

$$\mathbf{P} \triangleq \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}, \quad (3.22)$$

где элементы матрицы должны удовлетворять условиям:

$$p_{ik} \geq 0, \quad \sum_{k=1}^N p_{ik} = 1, \quad i = 1, 2, \dots, N. \quad (3.23)$$

На основе такого статистического описания, спектральная плотность мощности сигнала определяется следующим образом [28]:

$$S(f) = \frac{1}{T_s^2} \sum_{n=-\infty}^{\infty} \left| \sum_{i=1}^N p_i S_i \left(\frac{n}{T_s} \right) \right|^2 \delta \left(f - \frac{n}{T_s} \right) + \frac{1}{T_s} \sum_{i=1}^N p_i |S_i(f)|^2 + \frac{2}{T_s} \operatorname{Re} \left[\sum_{i=1}^N \sum_{k=1}^N p_i S_i^*(f) S_k(f) p_{ik} (e^{-j2\pi f T_s}) \right], \quad (3.24)$$

где S_i является преобразованием Фурье W сигнала s_i , которое заданно формулой (3.5), S_i^* – комплексное сопряжение S_i , а вероятности перехода

$$p_{ik}(z) \triangleq \sum_{n=1}^{\infty} p_{ik}^{(n)} z^n. \quad (3.25)$$

Величина $p_{ik}(n)$ определена как вероятность того, что элемент s_k передан через n интервалов после s_i . Это ik -тый элемент матрицы \mathbf{P}^n . Также, по определению, $p_{ik}^{(1)} \triangleq p_{ik}$.

Для сигналов, кодированных линейными кодами, которые не требуют памяти, последовательность элементов совершенно случайна, и последующий элемент не имеет никакой зависимости от настоящего элемента. В этом случае, $\mathbf{P}^n = \mathbf{P}$ для всех $n \geq 1$ и формула (3/24) упрощается:

$$S(f) = \frac{1}{T_s^2} \sum_{n=-\infty}^{\infty} \left| \sum_{i=1}^N p_i S_i \left(\frac{n}{T_s} \right) \right|^2 \delta \left(f - \frac{n}{T_s} \right) + \frac{1}{T_s} \sum_{i=1}^N p_i (1-p_i) |S_i(f)|^2 - \frac{2}{T_s} \sum_{i=1}^N \sum_{\substack{k=1 \\ i \neq k \\ i < k}}^N p_i p_k \operatorname{Re} [S_i(f) S_k^*(f)] \quad (3.26)$$

Далее, когда $N = 2$, а $p_1 = p$ и $p_2 = 1 - p$, СПМ принимает вид:

$$S(f) = \frac{1}{T_s^2} \sum_{n=-\infty}^{\infty} \left| p S_1 \left(\frac{n}{T_s} \right) + (1-p) S_2 \left(\frac{n}{T_s} \right) \right|^2 \delta \left(f - \frac{n}{T_s} \right) + \frac{1}{T_s} p(1-p) |S_1(f) - S_2(f)|^2. \quad (3.27)$$

При помощи полученных соотношений мы вычислим спектральные плотности мощности униполярных кодов NRZ, RZ, Манчестера, Миллера, FM0 и ВИМ. Сначала мы рассмотрим коды без памяти NRZ, RZ, Манчестера, ВИМ и ФИМ, затем с памятью – Миллера и FM0.

3.4.1.1. Код NRZ

Код NRZ состоит из двух элементарных сигналов:

$$\begin{aligned} s_1(t) &= A, & 0 \leq t \leq T_s, \\ s_2(t) &= 0, & 0 \leq t \leq T_s. \end{aligned} \quad (3.28)$$

Преобразования Фурье сигналов кода:

$$\begin{aligned} W_1(f) &= A T_s e^{-j\pi f T_s} \frac{\sin^2(\pi f T_s)}{\pi f T_s}, \\ W_2(f) &= 0. \end{aligned} \quad (3.29)$$

Используя формулу (3.27) и нормируя по энергии $E_s = A^2 T_s$, получим спектральную плотность мощности кода NRZ:

$$\frac{S(f)}{E_s} = p(1-p) \frac{\sin^2(\pi f T_s)}{(\pi f T_s)^2}. \quad (3.30)$$

Когда двоичные сигналы равновероятны ($p = 1/2$), формула (3.30) приобретает вид:

$$\frac{S(f)}{E_s} = \frac{\sin^2(\pi f T_s)}{(\pi f T_s)^2}. \quad (3.31)$$

3.4.1.2. Код RZ.

Униполярный код RZ подобен коду NRZ за исключением того, что элемент двоичной единицы возвращается в ноль через половину битового периода. Два элемента кода следующие:

$$\begin{aligned} s_1(t) &= A, & 0 \leq t \leq T_s/2, \\ s_2(t) &= 0, & 0 \leq t \leq T_s. \end{aligned} \quad (3.32)$$

Преобразования Фурье этих сигналов:

$$\begin{aligned} W_1(f) &= \frac{AT_s}{2} e^{-j\pi f T_s/2} \frac{\sin(\pi f T_s/2)}{\pi f T_s/2}, \\ W_2(f) &= 0. \end{aligned} \quad (3.33)$$

Произведя замену переменных в формуле (3.27) и нормируя по энергии $E_s = A^2 T_s$, получим СПМ кода RZ:

$$\frac{S(f)}{E_s} = \frac{p^2}{4T_s} \delta(f) + \frac{p^2}{4T_s} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \left(\frac{2}{n\pi} \right)^2 \delta\left(f - \frac{n}{T_s}\right) + \frac{p(1-p)}{4} \frac{\sin^2(\pi f T_s/2)}{(\pi f T_s/2)^2}. \quad (3.34)$$

Когда двоичные сигналы равновероятны ($p = 1/2$), формула (3.34) принимает вид:

$$\frac{S(f)}{E_s} = \frac{1}{16T_s} \delta(f) + \frac{1}{16T_s} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \left(\frac{2}{n\pi} \right)^2 \delta\left(f - \frac{n}{T_s}\right) + \frac{1}{16} \frac{\sin^2(\pi f T_s/2)}{(\pi f T_s/2)^2}. \quad (3.35)$$

3.4.1.3. Код Манчестера

В униполярной форме два элемента кода Манчестера следующие:

$$\begin{aligned} s_1(t) &= A, & 0 \leq t \leq T_s/2, \\ s_2(t) &= A, & T_s/2 \leq t \leq T_s. \end{aligned} \quad (3.36)$$

Преобразования Фурье сигналов кода Манчестера:

$$W_1(f) = \frac{AT_s}{2} e^{-j\pi f T_s/2} \frac{\sin(\pi f T_s/2)}{\pi f T_s/2},$$

$$W_2(f) = \frac{AT_s}{2} e^{-j3\pi f T_s/2} \frac{\sin(\pi f T_s/2)}{\pi f T_s/2}.$$
(3.37)

Используя формулу (3.27) и нормируя по энергии $E_s = A^2 T_s$, получим спектральную плотность мощности кода Манчестера:

$$\frac{S(f)}{E_s} = \frac{\pi}{64T_s} \delta(f) + \frac{1-4p(1-p)}{16T_s} \sum_{\substack{n=-\infty \\ n \neq 0 \\ n=\text{нечетное}}}^{\infty} \left(\frac{1}{n\pi}\right)^2 \delta\left(f - \frac{n}{T_s}\right) +$$

$$+ \frac{1}{16T_s} \sum_{\substack{n=-\infty \\ n \neq 0 \\ n=\text{четное}}}^{\infty} \left(\frac{1}{n\pi}\right)^2 \delta\left(f - \frac{n}{T_s}\right) + p(1-p) \frac{\sin^4(\pi f T_s/2)}{(\pi f T_s/2)^2}.$$
(3.38)

При равновероятных сигналах ($p = 1/2$), формула (3.38) приобретает следующий вид:

$$\frac{S(f)}{E_s} = \frac{\pi}{64T_s} \delta(f) + \frac{1}{16T_s} \sum_{\substack{n=-\infty \\ n \neq 0 \\ n=\text{четное}}}^{\infty} \left(\frac{1}{n\pi}\right)^2 \delta\left(f - \frac{n}{T_s}\right) + \frac{1}{4} \frac{\sin^4(\pi f T_s/2)}{(\pi f T_s/2)^2}.$$
(3.39)

Нормированные спектральные плотности мощности кодов NRZ, RZ и кода Манчестера представлены на рис.3.7.

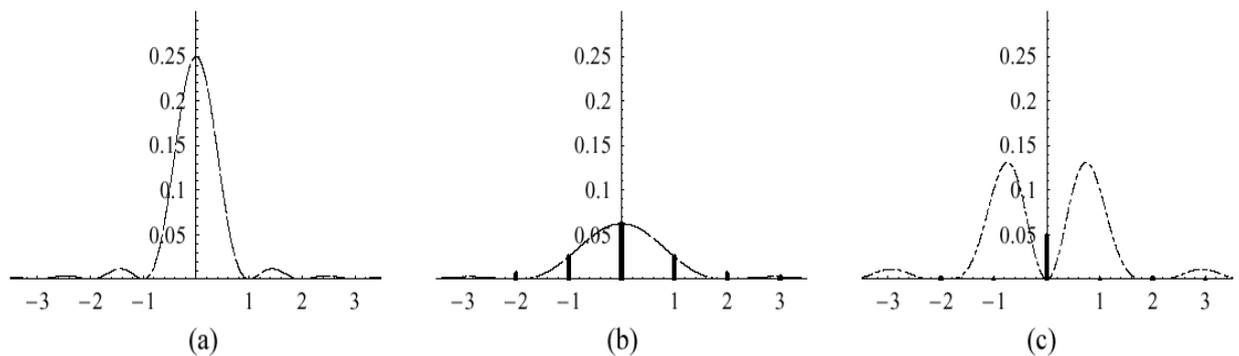


Рис.3.7. Нормированная спектральная плотность мощности в зависимости от $1/T$ для униполярных кодов (a) NRZ, (b) RZ, (c) Манчестера.

3.4.1.4. Коды ШИМ и ВИМ

До сих пор любой из видов линейного кодирования определял равными большую и малую длительности импульсов. Однако в зависимости

от вида связи в системе RFID (от считывателя к метке или от метки к считывателю) желательно оптимизировать энергию кодированного сигнала. Сигнал от считывателя должен максимизировать энергию, а сигнал от метки должен ее минимизировать. Поскольку линейные коды используют равные длительности импульсов, мощность сигнала не оптимальна для питания метки. Методы кодирования ШИМ и ФИМ располагают большей гибкостью. Будем рассматривать случай, когда энергия сигнала достаточно велика.

Благодаря гибкости, реализуемой при кодировании ШИМ и ФИМ, в зависимости от применения их определение может быть различным. Мы рассмотрим импульс, закодированный так, как показано на рис.3.6. Такая форма отличается от традиционного использования ее в других случаях, однако эти формы часто используются в системах RFID. Мы вычислим спектральную плотность мощности только для кода ШИМ, поскольку СПМ кодов ФИМ и ШИМ совпадают.

Найдем спектральную плотность мощности ШИМ модулированного по длительности сигнала с длительностью импульса $T_1 = r_1 T_s$, представляющего логическую 1, и $T_2 = r_2 T_s$, представляющего логический 0. Сигнал состоит из двух отдельных сигналов s_1 и s_2 с вероятностями появления $p_1 = p$ и $p_2 = 1 - p$ соответственно:

$$\begin{aligned} s_1(t) &= \begin{cases} A, & 0 \leq t \leq T_1, \\ 0, & T_1 \leq t \leq T_s, \end{cases} \\ s_2(t) &= \begin{cases} A, & 0 \leq t \leq T_2, \\ 0, & T_2 \leq t \leq T_s. \end{cases} \end{aligned} \quad (3.40)$$

Используя преобразование Фурье прямоугольных импульсов, получим:

$$\begin{aligned} W_1(f) &= Ar_1 T_s e^{-j\pi f r_1 T} \frac{\sin(\pi f r_1 T)}{\pi f r_1 T}, \\ W_2(f) &= Ar_2 T_s e^{-j\pi f r_2 T} \frac{\sin(\pi f r_2 T)}{\pi f r_2 T}. \end{aligned} \quad (3.41)$$

Подставляя в формулу (3.27) и нормируя по энергии $E_s = A^2 T_s$, получим:

$$\begin{aligned} \frac{S(f)}{E_s} &= \frac{\pi[p(r_1 - r_2) + r_2]^2}{2T_s} \delta(f) + \\ &+ \frac{1}{T_s} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1 - p(1-p) - p \cos(2\pi n r_1) + p(1-p) \cos[2\pi n(r_1 - r_2)] - (1-p) \cos(2\pi n r_2)}{2(n\pi)^2} \delta\left(f - \frac{n}{T_s}\right) + \\ &+ p(1-p) \frac{\sin^2[\pi f T_s (r_1 - r_2)]}{(\pi f T_s)^2}. \end{aligned} \quad (3.42)$$

Когда двоичные сигналы равновероятны ($p = 1/2$), формула (3.42) принимает вид:

$$\begin{aligned} \frac{S(f)}{E_s} &= \frac{\pi \left[\frac{1}{2}(r_1 - r_2) + r_2 \right]^2}{2T_s} \delta(f) + \\ &+ \frac{1}{T_s} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{3 - 2 \cos(2\pi n r_1) + \cos[2\pi n(r_1 - r_2)] - 2 \cos(2\pi n r_2)}{8(n\pi)^2} \delta\left(f - \frac{n}{T_s}\right) + \\ &+ \frac{1}{4} \frac{\sin^2[\pi f T_s (r_1 - r_2)]}{(\pi f T_s)^2}. \end{aligned} \quad (3.43)$$

С учетом предположения, что $r_1 = 1/2$ и $r_2 = 7/8$, формула (3.43) приобретает следующий вид:

$$\begin{aligned} \frac{S(f)}{E_s} &= \frac{121\pi}{256T_s} \delta(f) + \\ &+ \frac{1}{T_s} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{3 + \cos(3\pi n/4) - 2 \cos(\pi n) - 2 \cos(7\pi n/4)}{8(n\pi)^2} \delta\left(f - \frac{n}{T_s}\right) + \\ &+ \frac{1}{4} \frac{\sin^2[3\pi f T_s / 8]}{(\pi f T_s)^2}. \end{aligned} \quad (3.44)$$

На рис.3.8 показан вид СПМ, соответствующий расчету по формуле (3.44).

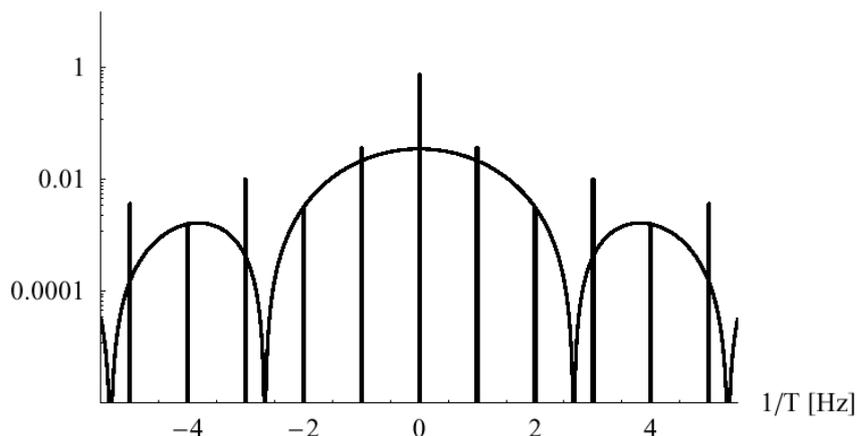


Рис. 3.8. Нормированная СПМ ШИМ сигнала, модулированного длительностями импульса, равными $7/8$ и $1/2$ битового периода T .

Из рис. 3.8 видно, что для ШИМ сигнала ширина спектра на уровне -20 дБ соответствует 3-ей спектральной линии и приблизительно равна $6/T$.

3.1.4.5. Коды Миллера и FM0

Вычисление спектральной плотности мощности кодов Миллера и FM0 более сложно, так как они являются кодами с памятью. Поскольку каждый последующий элемент этих кодов формируется на основе элемента, который анализируется в настоящее время, следует использовать формулу (3.24). Мы представим вычисление кода Миллера, приведенное в [27] и [29]. Спектральные плотности мощности для униполярных кодов Миллера и FM0 одинаковы.

На рис. 3.9 показаны четыре элементарных символа и четыре перехода кода Миллера.

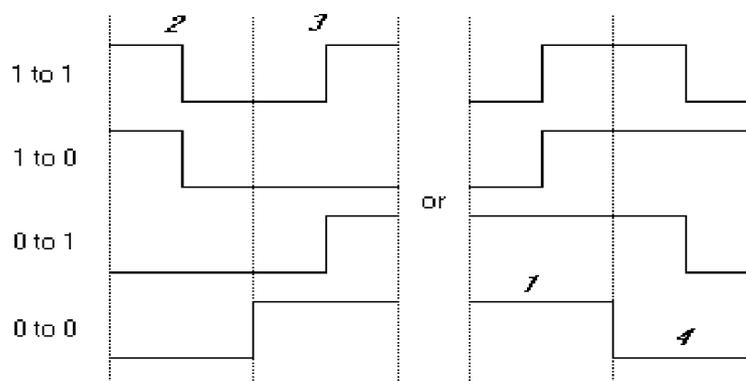


Рис.3.9. Переходы кода Миллера и его элементы, обозначенные 1, 2, 3 и 4.

Эти элементы можно задать математически:

$$\begin{aligned}
s_1 = -s_4 = A, \quad 0 \leq t \leq T_s, \\
s_2 = -s_3 = \begin{cases} A, & 0 \leq t \leq T_s/2, \\ -A, & T_s/2 \leq t \leq T_s. \end{cases}
\end{aligned} \tag{3.45}$$

Для того чтобы найти решение (3.24), необходимо найти матрицу вероятностей перехода, подобно формуле (3.22). Из рис. 3.9 видно, что каждый из элементов кода имеет вероятность появления, равную $1/4$, причем возможны только определенные переходы. Следовательно, матрицу вероятностей перехода можно представить следующим образом:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}. \tag{3.46}$$

Чтобы упростить бесконечный ряд, данный в формуле (3.25), мы найдем рекурсивные отношения, основанные на функции автокорреляции. Функция автокорреляции для дискретного набора сигналов кода Миллера следующая:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}. \tag{3.47}$$

Корреляционная матрица сигнала:

$$s_{ik} = \frac{1}{T_s} \int_0^{T_s} s_i(t) s_k(t) dt, \quad i, k = 1, 2, 3, 4. \tag{3.48}$$

Определим рекурсивные соотношения:

$$P^{4+l} S = -\frac{1}{4} P^l S, \quad l \geq 0. \tag{3.49}$$

Упростив при помощи этих соотношений формулу (3.24), получим нормированную спектральную плотность мощности:

$$\frac{S(f)}{E_s} = \frac{1}{2\theta^2(17+8\cos 8\theta)} \left(\begin{array}{l} 23 - 2\cos\theta - 22\cos 2\theta - 12\cos 3\theta + 5\cos 4\theta + \\ + 12\cos 5\theta + 2\cos 6\theta - 8\cos 7\theta + 2\cos 8\theta \end{array} \right). \quad (3.50)$$

Нормированная СПМ кода Миллера изображена на рис.3.10.

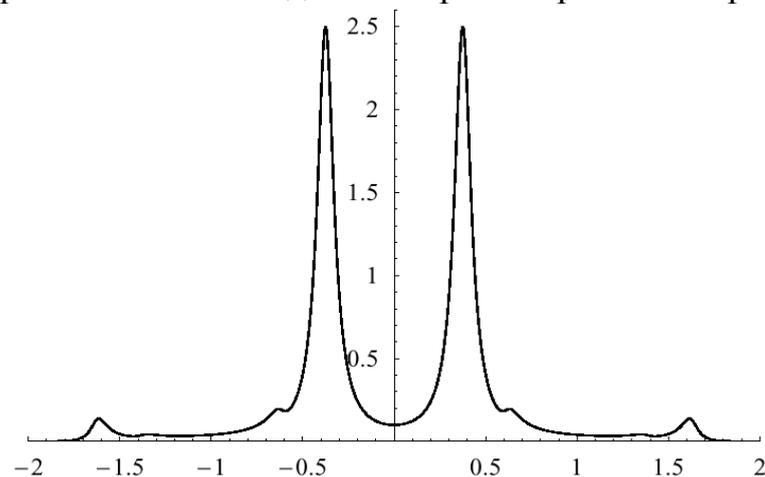


Рис. 3.10. Нормированная спектральная плотность мощности биполярного кода Миллера в зависимости от $1/T$.

3.5. Модуляция

Передача информации на несущей частоте f_c при помощи изменений амплитуды, частоты или фазы, или всех параметров одновременно, называется модуляцией [24]. В системах RFID информация передается по беспроводному каналу. Существует множество аналоговых и цифровых методов модуляции, но все они разными способами изменяют амплитуду, частоту или фазу несущего колебания. Далее мы будем рассматривать только двоичную цифровую модуляцию, называемую манипуляцией.

Существует три основных метода манипуляции:

- амплитудная манипуляция (ASK),
- двоичная фазовая манипуляция (BPSK),
- частотная манипуляция (FSK).

Во всех трех методах амплитуда, фаза или частота изменяется в соответствии с информацией, которую несет униполярный двоичный сигнал, представляющий цифровые данные в соответствии с выбранным способом кодирования. При амплитудной манипуляции несущий сигнал модулируется двумя амплитудами. При двоичной фазовой манипуляции (BPSK) несущий сигнал может иметь фазу 0 и 180 градусов. При частотной манипуляции (FSK) сигнал может иметь две несущих частоты.

В системе RFID информация передается от считывателя к метке и от метки к считывателю. Так как физические ограничения в системе различны, поэтому в этих двух случаях используются различные методы модуляции.

Считыватель должен генерировать такой сигнал, который доминировал бы над шумом и был способен обеспечить функционирование метки. Поскольку метка резко ограничена стоимостью и размерами (она содержит простейший приемник), сигнал, модулируемый считывателем, должен быть простым. Поэтому при передаче сигнала от считывателя к метке используется модуляция ASK, так как она предполагает относительно простое обнаружение. Обнаружение может быть когерентным, когда учитывается фаза несущего сигнала, или некогерентным, когда фаза сигнала не учитывается. Некогерентное обнаружение реализуется проще и дешевле. Детектирование огибающей – разновидность некогерентного обнаружения, когда детектируется огибающая амплитуды модулированного сигнала. Амплитудную модуляцию и детектирование огибающей мы рассмотрим далее в разделе 3.5.1.

Технические характеристики метки существенно отличаются от характеристик считывателя. Поскольку пассивная метка не имеет передатчика, она передает свою информацию при помощи вариации нагрузки и модуляции обратного рассеяния, которые основаны на принципах, рассмотренных в главе 2. Эти методы модуляции позволяют изменять амплитуду и фазу в зависимости от их реализации. Ввиду того, что по сравнению с сигналом считывателя мощность модулируемого меткой сигнала мала, часто для выделения этого сигнала применяют перенос спектра в сторону от центральной частоты при помощи поднесущей частоты. Модуляция нагрузки и обратного рассеяния с использованием поднесущей позволяют осуществить способы модуляции FSK и BPSK. Из-за особенностей и существенных различий в линиях связи от считывателя к метке и от метки к считывателю рассмотрим эти линии отдельно.

3.5.1. Кодирование и модуляция в прямой линии

В линии связи от считывателя к метке (ее часто называют прямой линией связи) должны быть выполнены следующие условия:

- метка должна получать достаточную для чипа энергию,
- метка должна обнаруживать сигнал,
- сигнал метки должен удовлетворять ограничениям по напряженности поля и полосе пропускания.

Первое условие очевидно. В пассивной системе RFID метка должна получать от считывателя необходимую энергию, так как если схемы кодирования и модуляции не будут иметь достаточного энергопитания, то метка не сможет функционировать. Большое значение имеет обнаружение

сигнала. Приемник метки должен быть максимально простым и, поэтому, он имеет невысокую чувствительность. Сложные приемники более чувствительны, но имеют более высокую стоимость. Кроме того, метка должна обеспечивать синхронизацию сигнала. Схемы кодирования предусматривают синхронизацию, либо задним фронтом – для ШИМ сигнала, либо переходом – для сигнала, кодированного Манчестерским кодом. Наконец, вероятность передачи ошибки должна быть минимизирована. Если достоверность данных не высока, связь в системе не будет обеспечена вне зависимости от оптимизации скорости передачи данных. Для удовлетворения регламентным ограничениям кроме схем кодирования и модуляции должны присутствовать схемы формирования и фильтрации сигнала, которые также предстоит рассмотреть.

В системах RFID низкой стоимости наиболее часто применяют амплитудную манипуляцию, а в метках применяют самые простые и дешевые приемники.

3.5.1.1. Амплитудная манипуляция

При двоичной амплитудной модуляции ASK униполярная информация передается двумя амплитудами несущего сигнала. Если меньшая из двух амплитуд равна нулю, то этот случай называется переключательной модуляцией ООК (On-Off shift Keying). В случае рассмотрения двух амплитуд – A_c и A_{\min} , определим индекс модуляции:

$$u = 1 - A_{\min} / A_c. \quad (3.51)$$

Когда $A_{\min} = A_c$, индекс модуляции u равен нулю и модуляция отсутствует. В случае, когда $A_{\min} = 0$, индекс $u = 1$, что соответствует 100 % модуляции амплитуды, то есть модуляции ООК. Мы можем описать ASK модулированный сигнал следующим образом:

$$v(t) = A_c[1 + u(s(t) - 1)]\cos(\omega_c t), \quad (3.52)$$

где $s(t)$ – униполярный двоичный сигнал (1 или 0), а ω_c – несущая частота в радианах в секунду. Используя формулы (3.18) – (3.20), мы можем найти спектр и спектральную плотность мощности модулированного сигнала. Модулирующий сигнал $g(t)$, основанный на кодовом сигнале $s(t)$, имеет следующий вид:

$$g(t) = A_c[1 + u(s(t) - 1)] + A_c(1 - u). \quad (3.53)$$

Спектр этого сигнала:

$$G(f) = u A_c S(f) + A_c(1 - u) \delta(f), \quad (3.54)$$

а спектральная плотность мощности:

$$G_g(f) = (u A_c)^2 S_s(f) + [A_c(1 - u)]^2 \delta(f). \quad (3.55)$$

Подставив последнюю формулу в (3.20), получим СПМ модулированного сигнала в области положительных частот:

$$S_v(f) = 1/4[(u A_c)^2 S_s(f - f_c) + [A_c(1 - u)]^2 \delta(f - f_c)], \quad f \gg 0. \quad (3.56)$$

Обратим внимание на то, что форма спектра модулированного сигнала $g(t)$ определяется спектром модулирующего сигнала $s(t)$, а его величина зависит от амплитуды несущего сигнала A_c и индекса модуляции u .

При индексе модуляции, равном 1 (в случае ООК), СПМ модулированного сигнала повторяет спектральную плотность мощности модулирующего сигнала, умноженную на $1/4$ квадрата амплитуды несущего сигнала. При этом отсутствует спектральная линия несущей частоты f_c . Уменьшение индекса модуляции приводит к уменьшению величины полос боковых частот, но увеличивает вес дельта функции на несущей частоте.

В системах RFID, особенно в частотных диапазонах с наиболее строгими ограничениями на полосу пропускания, обычно используется индекс модуляции меньше единицы. Отрицательными последствиями этого является уменьшение различия между высокими и низкими уровнями сигнала и увеличение вероятности ошибки.

3.5.1.2. Детектирование огибающей

Детектирование огибающей радиосигнала – это вид детектирования, часто используемая для обнаружения амплитудно-модулированных сигналов. По существу, выделяется и обрабатывается положительная часть огибающей реальной составляющей входного сигнала. На рис.3.11 изображена упрощенная схема детектора огибающей.

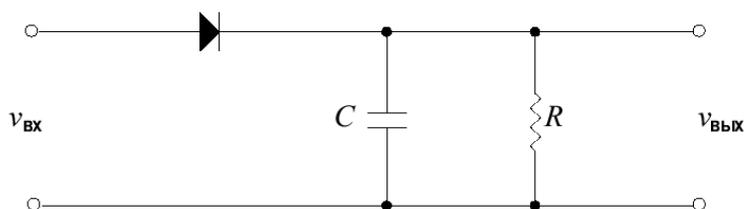


Рис.3.11. Детектор огибающей.

Резистор и конденсатор образуют фильтр низких частот, постоянная времени которого RC выбрана таким образом, что выходной сигнал отслеживает огибающую входного сигнала. Верхняя граница полосы пропускания низкочастотного фильтра должна быть гораздо меньше частоты несущего сигнала, но гораздо больше ширины полосы B входного модулированного сигнала:

$$B \ll 1/2\pi RC \ll f_c. \quad (3.57)$$

Выходной сигнал детектора огибающей является по существу огибающей входного сигнала, умноженной на коэффициент пропорциональности K . Для огибающей сигнала АСК из формулы (3.53) выходное напряжение определится следующим образом:

$$v_{\text{вых}} = K|g(t)| = KuA_c(1 - u). \quad (3.58)$$

Следует отметить, что при индексах модуляции $u < 1$ в выходном сигнале всегда присутствует постоянная составляющая.

3.5.2. Кодирование и модуляция в обратной линии

Условия и требования к линии связи от метки к считывателю (ее часто называют обратной линией связи) отличаются от условий и требований к прямой линии связи. Связь от метки к считывателю осуществляется при помощи вариации нагрузки и обратного рассеяния. В зависимости от варианта реализации связи модулируется и амплитуда, и фаза сигнала метки, или только амплитуда. При этом схемы кодирования и модуляции в метке должны иметь минимальное энергопотребление и ширину спектра сигнала. Так как уровни ответного сигнала метки невелики, они регламентируются не так строго, как сигналы большого уровня мощности считывателя. Однако следствием малого уровня сигнала является сложность его приема считывателем. Чаще всего излучаемый сигнал, который поступает на вход приемника считывателя, значительно превосходит сигнал, приходящий с метки, поэтому его обнаружение связано со значительными трудностями. Как правило эту проблему решают сдвигом сигнала на поднесущую частоту. В случае применения BPSK или FSK модуляции поднесущая легко реализуется посредством АМ и ФМ, которые используются при модуляции нагрузки и обратного рассеяния.

Теперь кратко рассмотрим аспекты использования поднесущих, а затем обсудим вопросы кодирования линии связи от метки к считывателю.

3.5.2.1. Метод поднесущей частоты

Метод поднесущей частоты реализуется при помощи модуляции нагрузки или модуляции обратного рассеяния на более высоких частотах, чем скорость передачи данных. Изменение фазы или частоты при модуляции поднесущей осуществляется соответственно BPSK или FSK. Некоторые системы RFID, которые, например, работают на частоте 13,56 МГц, используют частоту поднесущей 212 кГц. При этом благодаря модуляции в метке, боковые полосы сигнала располагаются на 212 кГц выше и ниже несущей частоты 13,56 МГц. Перенос сигнала на поднесущие частоты приводит к разнесению боковых полос сигнала метки и сигнала считывателя, что обеспечивает снижение уровня шумов и, соответственно повышает чувствительность приемника считывателя. Такое решение, однако, приводит к повышению рабочей частоты схемы синхронизации или генератора в метке и, следовательно, к увеличению ее энергопотребления.

3.5.2.2. Кодирование

Кодирование в обратной линии связи (от метки к считывателю) отличается от кодирования прямой линии связи (от считывателя к метке) тем, что энергия сигнала должна быть минимизирована. Уменьшение энергии отраженного сигнала позволяет обеспечить большее количество энергии, доступной для питания электронной схемы метки. Амплитуда сигнала, однако, должна быть достаточной для того, чтобы считыватель смог обнаружить сигнал. Обычно используются коды Манчестера, FM0 или другие. Сложность использования кода Манчестера состоит в том, что вариация фазы сигнала может приводить к ошибке в приеме информации считывателем из-за неоднозначности определения сдвига фазы на 180 градусов. С другой стороны, недостатки других методов – методов кодирования с переходами внутри информационного бита, состоят в том, что они требуют более высокочастотной синхронизации или меньшей скорости передачи данных.

3.5.3. Модуляция с распределенным спектром

При решении проблем связи, таких, например, как множественный доступ, защита от помех и подавление интерференций, учитывается ширина спектра, энергия сигнала и другие параметры [24]. Модуляция с распределенным спектром (Spread Spectrum – SS), которая осуществляется за счет распределения энергии сигнала по полосе частот намного большей, чем требуется собственно сигналу, позволяет осуществить множественный доступ, защиту от помех и подавление интерференций. Системы RFID обычно используют методы модуляции с распределенным спектром в UHF и

микроволновом диапазоне частот. На этих частотах, как правило, регламенты допускают для систем с распределенным спектром большие выходные мощности сигналов, по сравнению с узкополосными системами. Энергия распределяется в широкой полосе частот, уменьшая вероятность вредного влияния одних устройств на другие.

Известны два наиболее распространенных типа модуляции с распределенным спектром: прямая последовательность (Direct Sequence – DS) и скачкообразная перестройка частоты (Frequency Hopping – FH).

В DS модуляции псевдошумовой кодовый сигнал с периодом меньшим, чем битовый период информации, смешивается с сигналом сообщения, создавая выходной широкополосный сигнал. Чтобы восстановить сообщение, приемник должен демодулировать код той же псевдошумовой последовательностью, которая использовалась при передаче сигнала.

В системах FH со скачкообразной перестройкой частоты для передачи информации используют множество узкополосных каналов на частотах, выбранных из псевдослучайного списка, и перестраиваются на них. Это приводит к эффекту усреднения энергии в полосе частот перестройки. При этом для того, чтобы принять модулируемый сигнал приемники таких систем должны использовать тот же псевдослучайный список частот.

В пассивных системах RFID низкой стоимости сигналы с распределенным спектром часто используются иначе, чем в технологии связи. Для достижения полной DS или FH модуляции метке потребовалась бы чрезмерно сложная и энергоемкая обработка сигнала. Вместо этого, антенну метки проектируют таким образом, чтобы ее полоса пропускания охватывала всю ширину спектра сигналов, излучаемых считывателем. В этом случае метка получает всю мощность, заключенную в узкополосных перестраиваемых каналах или распределенную по полной ширине полосы частот при DS модуляции. Независимо от того, какая модуляция используется – DS или FH, сигналы считывателя модулируются по амплитуде и затем отражаются меткой. В технологии RFID такое решение достаточно эффективно, поскольку позволяет добиться более высокой мощности отраженного сигнала, как для узкополосных систем, так и для систем с распределенным спектром.

3.6. Вероятность ошибки

Вообще, есть два важных фактора для оценки системы связи – полоса пропускания и характеристики приема сигнала в присутствии шума [24]. Мы уже рассмотрели вопросы, связанные с полосой пропускания, теперь мы рассмотрим качество функционирования системы в присутствии шума.

Числовой оценкой качества функционирования аналоговых систем является отношение мощности сигнала к мощности шума. В цифровых системах критерием качества является вероятность ошибки в одном двоичном разряде или вероятность появления ошибочных битов (bit error rate – BER).

BER – это вероятность возникновения ошибки, когда система должна принять решение о приеме одного из двух возможных сигналов. В случае модуляции ASK это сигналы высокого или низкого уровня (в системах с ООК низкий уровень равен нулю). Приемник имеет порог принятия решения, выше которого сигналы считаются сигналами высокого уровня, и ниже которого – низкого уровня. Полная вероятность ошибки (BER) является суммой вероятностей ошибки, связанной с обоими сигналами.

Когда детектирование когерентное, используется большее количество информации – информация об амплитуде и фазе. Когда детектирование некогерентное, используется информация только об амплитуде. При этом ошибки функционирования системы увеличиваются. Тем не менее из-за простоты и невысокой стоимости в аппаратуре RFID обычно используют некогерентное детектирование огибающей.

В сигналах, модулированных ASK и ООК, присутствуют только два уровня сигналов. Каким образом эти уровни представляют отдельный бит, зависит от используемого вида кодирования. Однако, независимо от вида кодирования, если ошибки сделаны в процессе определения уровня сигнала, возникнет битовая ошибка. Некоторые виды кодирования могут обнаружить такую ошибку, а другие не могут. В таком случае могут быть использованы другие методы обнаружения и исправления ошибки.

В беспроводных каналах шумы могут иметь разнообразную природу. Обычно рассматривают два вида шума - это импульсные помехи и гауссовский шум. Гауссовский шум обычно является результатом фонового излучения, тепловых помех и дробового шума. В устройствах связи ближнего действия, особенно в пассивных системах RFID невысокой стоимости, обычно наиболее важны импульсные помехи, которые являются результатом интерференции многих других мешающих излучений. Для вычисления BER систем RFID рассмотрим как импульсные помехи, так и гауссовский шум.

3.6.1. Импульсные помехи

Источниками импульсных помех может быть интерференция сигналов посторонних источников или эффект многолучевого распространения. Эти помехи наиболее вредны для систем RFID, в которых используются сигналы малой мощности и простые детекторы. Гильбертова [30] модель ошибок, создаваемых импульсными помехами, представляется в виде цепи Маркова с двумя состояниями, при этом одно состояние может быть благоприятным,

когда вероятность ошибки невелика, а другое неблагоприятным, когда ошибки случаются с высокой вероятностью. Такую модель предложил Эллиотт [31], а другие авторы [32, 33] использовали и развили ее. Импульсные помехи могут искажать как отдельные биты, так и битовые блоки сигнала.

3.6.2. Гауссовский шум

При рассмотрении аддитивного гауссовского шума будем иметь в виду два уровня сигнала: с амплитудой, отличной от нуля, и с нулевой амплитудой. В зависимости от этих двух уровней плотности вероятности сигнала в присутствии шума различны. В общем случае плотность вероятности сигнала с шумом зависит от уровня сигнала. Наиболее простой случай – когда уровень сигнала равен нулю и шум является единственным компонентом приемника. Такая ситуация характерна ООК.

Однако в ASK с индексом модуляции меньше 1 всегда присутствует смесь сигнала с шумом. Мы рассмотрим различия в функциях плотности вероятности при изменении соотношения уровней сигнала и шума.

Сначала мы рассмотрим простой случай, когда амплитуда сигнала равна нулю. В этом случае сигнал на входе с детектора огибающей представляет собой шум n . В предположении, что шум гауссовский, мы можем разложить его на синфазную и квадратурную составляющие:

$$c(t) = n(t) = x(t)\cos(\omega_c t) - y(t)\sin(\omega_c t), \quad (3.59)$$

где x и y – независимые гауссовские случайные переменные. Преобразуем переменные x и y из декартовой формы в полярную с компонентами амплитуды r и фазы θ . После преобразования составляющих шума в полярную форму определится совместная плотность вероятности компонент r и θ . При использовании детектора огибающей нас интересует только амплитуда. Поэтому, после интегрирования совместной плотности вероятности по фазе θ , мы найдем плотность вероятности и функцию распределения случайной амплитуды r . Найденная функция распределения называется рэлеевским распределением.

Теперь мы можем рассмотреть случай, когда амплитуда сигнала отлична от нуля, это может быть максимальная амплитуда или амплитуда с меньшим уровнем, когда индекс модуляции меньше единицы. Полезный модулированный сигнал может быть задан в виде (3.52). Тогда смесь сигнала с шумом c на входе детектора:

$$c(t) = v(t) + n(t) = A_c[1+u(s(t) - 1)]\cos(\omega_c t) + x(t)\cos(\omega_c t) - y(t)\sin(\omega_c t). \quad (3.60)$$

Аналогично предыдущему случаю преобразуем синфазную и квадратурную составляющие компонент в полярную форму. После чего, с учетом зависимости компонент, определим совместную плотность вероятности r и θ . Затем, используя модифицированную функцию Бесселя первого порядка, найдем интегральную функцию распределения [34].

Результирующая функция распределения смеси сигнала с шумом называется распределением Райса. Следует обратить внимание, что при амплитуде сигнала равной нулю плотность вероятности описывается рэлеевским распределением, а при амплитуде намного большей, чем интенсивность шума – гауссовским распределением.

Учитывая вероятности случайных уровней сигнала, мы можем найти совокупную вероятность ошибки, которая, как было отмечено выше, соответствует сумме областей плотности вероятности, каждая из которых представляет собой ошибку одного из двух уровней сигнала. Следовательно,

$$P_e = P_1 \int_{-\infty}^b f(r|v_1) dr + P_2 \int_b^{\infty} f(r|v_2) dr, \quad (3.61)$$

где P_1 и P_2 – вероятности появления 1-го и 2-го сигнала, соответственно, а параметр b – пороговый уровень. При модуляции ООК первая плотность вероятности соответствует рэлеевскому распределению, вторая – распределению Райса. В случае ASK с индексом модуляции меньше 1 обе плотности распределения описываются распределением Райса. К сожалению, для распределений Райса интегрирование функции Бесселя в аналитической форме не представляется возможным. Однако задача может быть решена численно. В результате можно сделать вывод: чем меньше различие в уровнях сигнала, тем больше вероятность ошибки. Кроме того, более низкие уровни сигнала относительно среднеквадратичного действующего значения шума дают большую BER. Чтобы получить минимальную вероятность ошибки, необходимо оптимизировать пороговый уровень b . Такая оптимизация, однако, требует знания соотношения сигнал/шум, что на практике реализовать крайне сложно.

При передаче сигналов от считывателя к метке и от метки к считывателю, из-за необходимости применения упрощенных детекторов, малых уровней сигнала и, в результате, относительно высоких BER, для повышения достоверности приема информации целесообразно использовать алгоритмы обнаружения и исправления ошибок. Рассмотрим такие методы.

3.7. Обнаружение и коррекция ошибок

Хотя излучаемый считывателем сигнал имеет достаточно высокий уровень, метка содержит упрощенные схемы детектирования и обработки

сигнала, что приводит к высокой вероятности появления ошибок в прямой линии связи. В обратной линии связи сигнал, переданный меткой, имеет относительно низкий уровень. При этом, несмотря на то, что считыватель RFID имеет достаточно чувствительный приемник и гораздо более сложную обработку сигнала, из-за малого уровня сигнала также возможна высокая вероятность появления ошибок. Следовательно, в обеих линиях связи целесообразно использовать процедуры обнаружения или исправления ошибок.

Операции кодирования, дополненные возможностью обнаружения и исправления ошибок, требуют добавления дополнительных битов к первоначальному сообщению. В зависимости от количества битов и способа, которым они добавлены, эти коды могут исправлять или только обнаруживать ошибки. В общем случае эффективные коды с обнаружением и исправлением ошибок требуют большого количества дополнительных битов, широкой полосы пропускания и сложной обработки. В системах RFID время передачи информации ограничено и достаточно строго регламентируется. Следовательно, целесообразно применение методов, которые позволяют только обнаруживать ошибки.

Типичные методы обнаружения ошибок включают проверку четности, продольный контроль по избыточности (Longitudinal Redundancy Checking – LRC) и циклический контроль по избыточности (Cycling Redundancy Checking – CRC) [6]. При проверке четности к строке битов добавляется дополнительный бит, который показывает: четное или нечетное количество бит содержится в строке. Такая проверка очень проста, но все же бесполезна, если в строке имеется четное количество ошибок. Продольный контроль по избыточности требует действия рекурсивного исключающего ИЛИ (XOR) для каждого байта в пределах блока данных. Результаты операции добавляются в конец строки и передаются. При приеме выполняется та же самая процедура. Если результат проверки не равен нулю, то произошла ошибка. При проверке четности многократные ошибки могут отменять друг друга. Циклический контроль по избыточности более надежен, но требует более сложной обработки. Операция CRC обеспечивает почти уникальную идентификацию битовой строки. Чем длиннее CRC, тем большее количество данных может быть надежно проверено. Поскольку циклы передачи данных в системах RFID относительно коротки, обычно используются 8-битный или 16-битный CRC.

3.8. Резюме

В этой главе мы рассмотрели основные принципы обеспечения связи в пассивных системах RFID низкой стоимости. Следует обратить внимание,

что при выборе способов кодирования и модуляции, как в прямой, так и в обратной линии связи, возможны различные решения.

В последующем мы увидим, что кодирование и модуляция влияют не только на ширину спектра сигнала и его спектральную плотность. В следующей главе мы рассмотрим, как административные регламенты ограничивают ширину спектра, а в дальнейшем – как это влияет на параметры и характеристики метки.

Глава 4

4. Электромагнитная совместимость и регламенты

4.1. Введение

В предыдущих двух главах мы рассмотрели необходимые основы электродинамики и связи, которые важны для понимания того, как функционируют системы RFID. В частности мы увидели, что рабочая частота, мощность и ширина спектра являются наиболее важными параметрами, которые определяют основные технические характеристики систем RFID – дальность, быстродействие и надежность передачи данных. Теперь мы рассмотрим вопросы электромагнитной совместимости и административные регламенты, которые задают диапазоны частот, мощность и ширину спектра сигнала, а также как эти регламенты влияют на аппаратную совместимость систем RFID во всем мире.

4.2. Необходимость распределения частот

При всем многообразии применения систем RFID сопровождение объектов в глобальной цепи доставки товаров от производителя к потребителю является одним из самых больших потенциальных рынков. При этом существенным является то, что метки должны функционировать в различных странах. Чтобы это было возможным, метки должны быть совместимыми со считывателями различных стран, а метки и считыватели одновременно должны удовлетворять требованиям по электромагнитной совместимости. Эти требования формулируются национальными администрациями и называются регламентами. Аппаратная совместимость систем RFID может быть обеспечена двумя способами: или любая метка должна функционировать во всех возможных диапазонах частот, или во всем мире должны использоваться общие частотные диапазоны.

Реализация меток, которые работают в очень широкой полосе частот (например, от 125 кГц до 13,56 МГц) требует значительных технических и стоимостных затрат. Некоторые антенны могут иметь удовлетворительные характеристики в пределах ограниченных диапазонов (например, в диапазоне 902 – 928 МГц в США) или в близко расположенных диапазонах (868 МГц в Европе и 915 МГц в США). Так как недорогие метки могут быть созданы только в сравнительно узком частотном диапазоне, то для обеспечения совместимости аппаратуры различных производителей, необходимо согласовывать вопросы выделения частотных диапазонов для систем RFID во всем мире.

4.3. Всемирная структура регулирующих организаций

В настоящее время в мире существует хорошо организованная структура, состоящая из международных, региональных и национальных регулирующих организаций. Кратко рассмотрим эти организации.

4.3.1. Организация ИТУ.

Международный союз ИТУ (International Telecommunication Union), в который входят все независимые государства [35], занимается распределением частот и публикацией таблиц их распределения (S5) [36]. Кроме значительного числа иных функций и целей, ИТУ для уменьшения влияния различных служб, которые используют радиочастотный спектр устанавливает взаимодействие и координирует усилия ее членов по выделению частотных диапазонов [ИТУ]. Отделение радиосвязи (ИТУ-R) отвечает за функционирование всех служб связи, использующих радиоволны. Хотя ИТУ стремится к единому размещению частот, в тех случаях, когда страны требуют использовать диапазоны, отличные от используемых во всем мире, делаются исключения, которые фиксируются в частотных таблицах. Для обозначения областей с особым частотным распределением ИТУ делит весь мир на регионы (S5.2) [36].

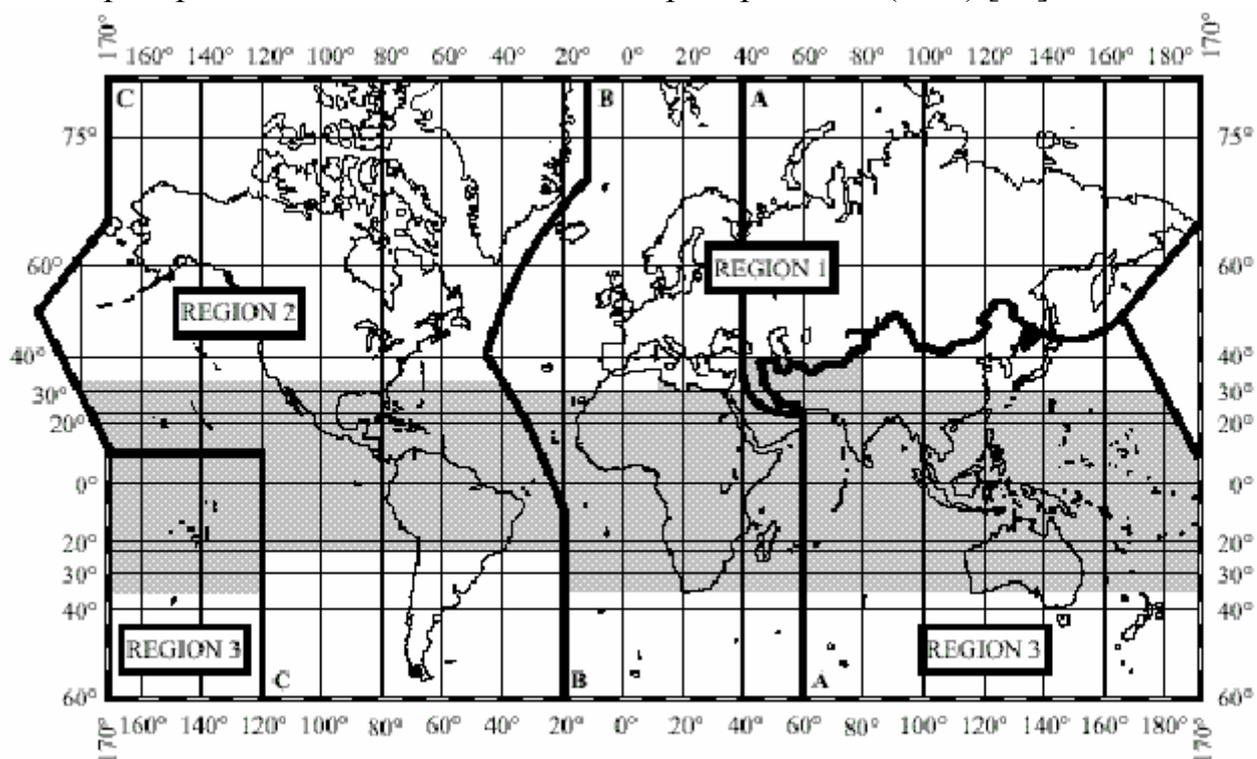


Рис. 4.1.1. Регионы ИТУ (S5.2) [36].
Region - Регион.

4.3.2. Региональные и национальные органы

Тогда как ИТУ развивает международную кооперацию и вырабатывает рекомендации для взаимодействия различных организаций, именно региональные и национальные органы устанавливают частотные диапазоны и предельные уровни излучения. Как большинство мировых экономик, так и потенциально наибольшие рынки технологии RFID, расположены в Европе, США и Японии. Поэтому мы акцентируем свое внимание на стандартах и соответствующих телекоммуникационных организациях этих регионов.

Европейская организация почтовых и телекоммуникационных организаций (European Conference of Postal and Telecommunications Administrations – СЕРТ) имеет дело с независимыми администрациями по вопросам почты и телекоммуникаций. Членами СЕРТ являются 43 страны центральной и восточной Европы. СЕРТ включает три комитета, первый – СЕРР (Comite europeen des regulateurs postaux) – по почтовым вопросам, и два других по вопросам телекоммуникаций – ЕРС (European Radiocommunication Committee) и ЕСТРА (European Committee for Regulatory Telecommunication Affairs). Эти комитеты одобряют гармонизирующие документы, рекомендации и решения, которые обычно готовятся их рабочими группами. Комитет ЕРС имеет в своем составе подразделение ЕРО (European Radiocommunications Office), которое проводит исследования и поддерживает активность комитетов [37].

Организация ETSI (European Telecommunications Standards Institute), учрежденная СЕРТ в 1988 г. (одновременно с СЕН и СЕНЕЛЕС), является одной из трех Европейских Организаций Стандартов (European Standard Organization – ЕSO). Она отвечает за разработку телекоммуникационных стандартов в Европе. 37 национальных организаций по стандартам из 34 стран, а также соответствующие организации России и Турции обеспечивают детальную проработку, утверждение и внедрение стандартов [38].

Организациями, ответственными за телекоммуникационные стандарты и регламенты в США и Японии, являются FCC (Federal Communications Commission) [39] и Министерство Почт и Телекоммуникаций (Ministry of Posts and Telecommunications – МРТ) [40] соответственно. В Японии уполномоченным центром по эффективному использованию радиочастотного спектра является АРИВ (Association of Radio Industries and Businesses). Кроме многих собственных задач АРИВ разрабатывает детальные национальные стандарты на основе менее подробных требований МРТ регламентов [41].

4.4. Диапазоны частот RFID

Для описания радиопередающих устройств, которые обеспечивают одностороннюю или двухстороннюю связь и которые обладают слабой возможностью создавать помехи другим радиотехническим устройствам СЕРТ использует термин Short Range Devices – SRD. В устройствах SRD антенны могут быть встроенными или внешними и могут использоваться любые виды модуляции [42]. ИТУ в рекомендациях 213/1 поддерживает это определение широкого класса радиоэлектронных приборов и, более того, рекомендует национальным администрациям, чтобы в определяемых ими параметрах при использовании устройств SRD отсутствовали излишне жесткие ограничения. Во всем мире прилагаются усилия по гармонизации использования радиоспектра в этом классе устройств.

Аппаратура RFID относится к классу устройств SRD и может функционировать только в назначенных частотных диапазонах, выделенных в большинстве стран. Рассмотрим эти диапазоны.

4.4.1. Диапазоны ISM

В таблицах распределения частот ИТУ выделяет определенные частоты для использования в промышленности, науке и медицине (Industrial, Science and Medicine – ISM). Эти диапазоны предназначены для применения в устройствах промышленного, научного и медицинского назначения, не предназначенных для дистанционного использования. Излучение этих устройств все же может оказывать помеховое влияние на работу других радио и телекоммуникационных служб. Поэтому для их использования выделяют определенные диапазоны частот [43]. Основными диапазонами ISM являются (RR S5.150) [36]:

- 13,553 - 13,567 МГц;
- 26,957 - 27,283 МГц;
- 40,66 – 40,70 МГц;
- 902 - 928 МГц (в Регионе 2);
- 2,4 - 2,5 ГГц;
- 5,725 - 5,785 ГГц;
- 24 - 24,25 ГГц.

По усмотрению национальных администраций в качестве диапазонов ISM могут выделяться дополнительные диапазоны (RR S5.138) [36]:

- 6,675 - 6,795 МГц;
- 61,0 - 61,5 ГГц;
- 122 - 123 ГГц;
- 244 - 246 ГГц.

В десяти странах Европы диапазон 433,05 - 434,79 МГц также является ISM диапазоном (S5.280) [36]. В других странах Региона 1 он доступен для использования при условиях, оговоренных RR S5.138. Диапазоны ISM также могут использоваться другими радиослужбами при условии, что они будут совместимыми с устройствами ISM. При этом ИТУ, не налагая ограничений на излучаемую мощность, рекомендует, чтобы все страны (администрации) добивались минимизации излучения в ISM диапазонах (S15.13) [36]. Одновременно в рамках резолюции RR № 63 между ИТУ и ИЕС обсуждаются возможности ограничения влияния помех в диапазонах ISM другим радиослужбам.

4.4.2. Другие диапазоны (LF, HF и UHF)

В большинстве регионов ИТУ разрешает без лицензий использовать в дополнение к диапазонам частот ISM низкочастотный диапазон LF (Low Frequency) 9 кГц – 135 кГц и HF (High Frequency) 13,56 МГц.

Согласно рекомендациям CEPT 70-03 со ссылками на соответствующие стандарты ETSI для использования в странах CEPT в разделе неспецифических устройств NS-SRD (Non Specific SRD) разрешены дополнительные UHF (Ultra High Frequency) диапазоны [44]. В частности это диапазон 868 МГц – 870 МГц. В настоящее время принимается решение по расширению этого диапазона до 865 МГц – 870 МГц.

Таблица 4.1. Диапазоны LF, HF, UHF и ISM, разрешенные в мире.

Частота	Диапазон частот	Ближняя/дальняя зона	$\lambda/2$	Тип	Регион	Примечания
9-135 кГц		382 м	1200 м			На поддиапазонах
6,78 МГц	± 15 кГц	7,04 м	22,1 м	ISM	Все*	
13,56 МГц	± 7 кГц	3,52 м	11,1 м	ISM	Все	
27,12 МГц	± 163 кГц	1,76 м	5,53 м	ISM	Все	
40,68 МГц	± 20 кГц	1,17 м	3,69 м	ISM	Все	
433,92 МГц	± 870 кГц	11,0 см	34,6 см	ISM	1**	
869,525 МГц	± 125 кГц***	5,5 см	17,3 см	NS-SRD	CEPT	Предлагается 865–870 МГц
915 МГц	± 13 МГц***	5,2 см	16,4 см	ISM	2	FH 250 /500 кГц
2,45 ГГц	± 50 МГц***	1,9 см	6,12 см	ISM	Все	FH в Европе и США
5,8 ГГц	± 75 МГц***	8,2 мм	2,59 см	ISM	Все	
24,125 ГГц	± 125 МГц	2,0 мм	6,2 мм	ISM	Все	

*На усмотрение администраций (RR S5.138) [36].

** Только в 10 Европейских странах (RR S5.280) [36]. Возможно использование в других странах региона 1, в соответствии с условиями RR S5.138 [36].

***Может использоваться скачкообразная перестройка частоты (FH) или широкополосные сигналы.

4.5. Напряженность поля, ширина спектра и методы измерения

До того как приступить к рассмотрению ограничений уровня излучения и полосы пропускания, мы остановимся на единицах измерений и методах измерений этих параметров.

Регламенты радиосвязи, устанавливаемые национальными администрациями, оперируют тремя предельными значениями параметров: мощности, напряженности поля и ширины спектра сигнала, а также методами измерения этих параметров. Мощность и напряженность поля могут быть определены в единицах излучаемой мощности или в единицах напряженности электрического или магнитного поля в ближней или дальней зоне. Ширина спектра сигнала может определяться на уровне 6 дБ, 20 дБ, 30 дБ или больше. Методы измерений параметров могут быть реализованы при помощи квазипикового детектора или детектора среднего уровня поля. Для того чтобы понять, как соотносить требования различных регламентов, необходимо привести их к общим методикам и единицам измерений. Для этого рассмотрим соответствующие формулы преобразования.

4.5.1. Преобразования напряженности поля и мощности

Для того чтобы сравнивать пределы напряженности поля и уровня мощности, устанавливаемые различными администрациями, необходимо преобразовать их к общим единицам измерения. Для большинства частот FCC устанавливает предельные уровни излучения в единицах излучаемой мощности или напряженности электрического поля в дальней зоне. В свою очередь, CEPT/ESTI устанавливает пределы в единицах измерения излучаемой мощности или в единицах измерения напряженности магнитного поля в ближней и дальней зоне. МРТ устанавливает пределы в единицах измерения излучаемой мощности или напряженности электрического поля в дальней зоне.

4.5.1.1. Дальняя зона

Как мы видели в главе 2, электрическое и магнитное поля в дальней зоне связаны друг с другом сопротивлением свободного пространства η_0 . Если принять, что максимальная напряженность электрического поля на расстоянии R_E равна E_{max}^{ff} , мы можем определить напряженность магнитного поля H_{max}^{ff} на другом расстоянии R_H . Для свободного пространства такие преобразования выглядят следующим образом:

$$E_{max}^{ff} = \frac{Idl}{4\pi R_E} \eta_0 k, \quad (4.1)$$

$$E_{\max}^{ff} = \frac{Idl}{4\pi R_H} k. \quad (4.2)$$

$$E_{\max}^{ff} = E_{\max}^{ff} \frac{R_E}{\eta_0 R_H}. \quad (4.3)$$

После нахождения напряженностей электрического и магнитного поля в какой-либо точке пространства дальней зоны, мы можем вычислить излучаемую мощность. Сначала найдем среднюю плотность мощности поля на расстоянии R_{EH} :

$$S = \frac{E_{ff}^2}{2\eta_0} = \frac{H_{ff}^2 \eta_0}{2} = \frac{P_t G_t}{4\pi R_{EH}^2}. \quad (4.4)$$

По определению EIRP является мощностью на входе антенны P_t , умноженной на ее усиление относительно изотропной антенны G_t :

$$EIRP = \frac{E_{ff}^2}{2\eta_0} 4\pi R_{EH}^2. \quad (4.5)$$

Используя выражение (2.48), теперь мы можем перейти к мощности ERP. Следует заметить, что, если E^{ff} и H^{ff} в формулах (4.4) и (4.5) выражены в среднеквадратических значениях, коэффициент 2 в этих формулах будет равен 1.

4.5.1.2. Ближняя зона

В ближней зоне взаимоотношения между электрическим и магнитным полем более сложны. Для проведения нужных преобразований следует учитывать тип антенны. Так как системы ближнего поля являются индуктивными или емкостными, необходимо рассмотреть оба типа антенн. Для индуктивных систем примем аппроксимацию антенны в виде малой рамки, а расстояние от антенны много большим радиуса рамки (уравнения (2.15) – (2.20)).

Если согласно уравнению (4.1) задать на некотором расстоянии R_E напряженность электрического поля в дальней зоне E_{\max}^{ff} , мы можем найти магнитное поле на расстоянии R_H в ближней зоне. Для этого сначала, используя уравнение (2.18), получим:

$$H_{\max}^{ff} = \frac{Idl}{4\pi k R_H^3} 2. \quad (4.6)$$

Теперь, комбинируя уравнения (4.1) и (4.6), получим оценочную зависимость между максимальным (пиковым) значением напряженности магнитного поля в ближней зоне с аналогичным значением электрического поля в дальней зоне:

$$H_{\max}^{nf} = E_{\max}^{ff} \frac{2R_E}{\eta_0 k^2 R_H^3}. \quad (4.7)$$

4.5.2. Ширина спектра

Если не оговаривается иначе, под шириной спектра понимается диапазон частот, в котором спектральные компоненты превышают некоторый заданный уровень. Администрации обычно оговаривают определенный уровень излучения, в котором могут присутствовать спектральные компоненты сигнала. При этом в некоторых диапазонах частот может допускаться только паразитное излучение.

Иногда ширина спектра обозначается как ширина спектра по определенному уровню n дБ, однако, в большинстве случаев она определяется путем сравнения предельного уровня основной напряженности поля или мощности с предельными уровнями боковых полос или прилежащего паразитного излучения. В этом случае ширина спектра по уровню n-дБ является отношением предельного уровня излучения на основной или центральной частоте к предельному уровню паразитного излучения или боковых полос, выраженном в децибелах.

Паразитное излучение обычно определяется как излучение, не включающее несущую частоту и боковые полосы. Оно часто используется в качестве предельного уровня внеполосного излучения модулированного сигнала [45].

4.5.3. Детектирование

Для того чтобы понимать, как регламенты влияют на характеристики систем RFID, крайне важно разобраться в методиках измерения. Большинство администраций в зависимости от несущей частоты и конкретного устройства проводят испытания оборудования или на открытом пространстве, или в безэховых камерах. Устройства испытываются во всех возможных угловых ориентациях и с изменением высоты для того, чтобы определить максимальную напряженность поля или мощность излучения на фиксированном расстоянии (FCC 15.31) [46]. Как правило, устройства испытываются в нормальном рабочем положении. Особый интерес представляет тип детектора, используемого при измерениях.

Обычно исследуют непрерывные или импульсные сигналы. Непрерывные сигналы имеют узкую ширину спектра. Импульсные сигналы, которые появляются в некоторые моменты времени с определенной частотой повторения, имеют более широкий спектр. При равных амплитудах сигналов обычно считается, что непрерывные сигналы более вредны.

Использование мгновенного пикового детектора для измерения уровня излучения не позволяет определить возможное вредное воздействие импульсных сигналов на окружающую среду. Поэтому в этом случае для оценки влияния излучения применяют усреднение сигнала во времени при помощи RC -фильтра (рис.4.2.) [47].

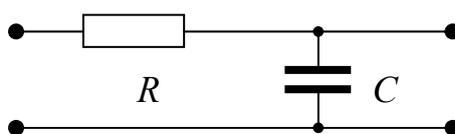


Рис.4.2. RC -фильтр.

Детекторы среднего значения предназначены для работы как в частотном диапазоне от 9 кГц до 1 ГГц [45], так и выше 1 ГГц [46]. Недостаток детекторов среднего значения состоит в том, что импульсы с малой скважностью имеют чрезвычайно малое среднее значение, что приводит к большим погрешностям измерения и, вследствие этого, к возможной недооценке степени влияния воздействия излучения на окружающую среду.

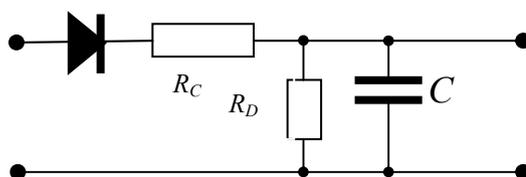


Рис.4.3. Квазипиковый детектор.

По этой причине обычно используют квазипиковые детекторы, которые состоят из зарядной и разрядной цепей (рис. 4.3). Такой детектор рассчитывается на определенную амплитуду, энергию импульса и частоту повторения сигнала. Чаще всего квазипиковые детекторы используются в диапазоне частот ниже 1 ГГц.

Исходные данные для проектирования и использования квазипиковых детекторов приводятся в публикации 16 ИЕС специального международного комитета по специальным воздействиям (CISPR) [48]. Технические характеристики детекторов (постоянная времени и полоса пропускания) различны в зависимости от диапазона частот. Всего определено четыре диапазона рабочих частот: 9 – 150 кГц (диапазон А), 150 кГц – 30 МГц

(диапазон В), 30 – 300 МГц (диапазон С) и 300 МГц – 1 ГГц (диапазон D). Выше 1 ГГц обычно используют спектральные анализаторы с пиковыми детекторами.

Для квазипикового детектора полоса пропускания по уровню 6 дБ в диапазоне А определена равной 220 Гц, в диапазоне В – 9 кГц, а в диапазонах С и D – 120 кГц. Если частота повторения импульсов превышает указанные частоты, детектор действует как детектор среднего значения и не требуется введения корректирующих коэффициентов. Радиотехнические системы с большой шириной спектра имеют частоты повторения, превышающие приведенные значения. Однако узкополосные системы с несущими частотами, например, менее 135 кГц, 6,78 и 13,56 МГц, могут иметь ширину спектра излучаемых сигналов менее 9 кГц или 220 Гц. В таких случаях вводится корректирующий коэффициент, связанный с частотой повторения импульсов [48]. Величина корректирующего коэффициента может быть существенной и превышать 10 дБ.

4.6. Частоты и ограничения

Поскольку системы RFID в ближней зоне используют электрические и магнитные поля, а в дальней – электромагнитные волны, рассмотрим их отдельно.

В таблице 4.1 для каждого диапазона приведены границы ближней зоны и размеры половины длины волны (для резонансных антенн). Очевидно разделение на системы ближней и дальней зон поля. Системы от 125 кГц до 40,6 МГц относятся к системам ближней зоны, так как граница их ближней зоны и полудлина волны достаточно велики. Системы с частотами от 433 МГц до 24 ГГц работают в дальней зоне, так как полудлина волны уменьшается до приемлемых размеров, а граница их ближней зоны становится небольшой.

Как правило, национальные регламентирующие органы в некоторых частотных диапазонах стремятся устанавливать общие для всех, согласованные предельные уровни паразитного излучения и других параметров. Мы рассмотрим и конкретные ограничения в интересующих нас диапазонах, и общие ограничения в окрестностях этих диапазонов.

Необходимо отметить, что в отдельных случаях, в промышленных высокочастотных устройствах, например, поля могут экранироваться, поэтому уровни излучения, превышающие допустимые нормы, существуют только внутри таких устройств. Что касается технологии RFID, мы имеем в виду неэкранированные системы.

Регламенты, на которые мы будем часто ссылаться, включают Европейские стандарты ETSI EN 300330 (Индуктивные рамки от 9 кГц до 25 МГц) [45], EN 300220 (Устройства SRD от 25 МГц до 1000 МГц) [49], EN

300440 (Устройства SRD от 1000 МГц до 40 ГГц) [50], рекомендации СЕРТ 70-03 (Устройства SRD во всех частотных диапазонах) [44], а также регламенты США FCC часть 15 [46] и Закон о радио (Radio Law) МРТ Японии.

4.6.1. Частоты RFID в ближней зоне

Каждая администрация устанавливает свои предельные уровни и ограничения на различных частотах. FCC определяет электрическое поле на некотором расстоянии в дальней зоне. ETSI/СЕРТ определяет магнитное поле на расстоянии 10 м, которое в зависимости от частоты находится или в ближней, или в дальней зоне поля. МРТ определяет электрическое поле на расстоянии 3 м на частоте около 16 МГц, при этом для других частот электрическое поле определяется по точке перехода от ближней зоны к дальней зоне. Все это будет предметом рассмотрения в следующих разделах.

4.6.1.1. Диапазон 9 кГц – 135 кГц

Хотя эти частоты и не являются диапазоном ISM, они часто используются для радиосвязных устройств небольшой дальности и ISM устройств. Кроме того, этот диапазон используют некоторые службы большой дальности действия, включая радионавигацию, службу времени и военные радиослужбы [ITU-R SM 1056/RR]. Нормы напряженности поля здесь обычно выше, чем в других диапазонах.

СЕРТ подразделяет весь диапазон на 5 поддиапазонов, в основном для обеспечения защиты средств радиосвязи от нежелательного взаимного влияния. Частоты от 9 кГц до 135 кГц также предназначены для индуктивных применений (рекомендации 70-03, приложение 9). Поддиапазоны и разрешенные уровни мощности в них приведены в табл. 4.2. [44,45].

Таблица 4.2. Частоты от 9 до 135 кГц, согласно СЕРТ/ETSI разрешенные для индуктивных устройств SRD.

Полоса частот	Напряженность магнитного поля	Детектор
9 – 59,75 кГц	72 дБмкА/м на 10м*	Квазипиковый
59,75 – 60,25 кГц	42 дБмкА/м на 10м**	Квазипиковый
60,25 – 70,0 кГц	72 дБмкА/м на 10м***	Квазипиковый
70 – 119 кГц	42 дБмкА/м на 10м	Квазипиковый
119 – 135 кГц	72 дБмкА/м на 10м***	Квазипиковый
135 кГц – 1 МГц	37,7 дБмкА/м на 10м***	Квазипиковый

* На 30 кГц спад 3,5 дБ/октаву (эквивалентно $72 - 3,5 \log_2(f[\text{кГц}]/30)$ дБмкА/м на 10 м).

** Для рамочных антенн площадью менее $0,05 \text{ м}^2$ предел на 10 дБ ниже.

Для рамочных антенн площадью от 0,05 до 0,16 м² предел равен приведенному значению + 10 log (площадь/0,16 м²).

*** На 135 кГц спад 3 дБ/октаву (эквивалентно $37,7 - 3 \log_2(f[\text{кГц}]/135)$ дБмкА/м на 10 м).

Регламент FCC в разделе 15.209 задает общие пределы излучения. Для частот между 9 кГц и 490 кГц напряженность электрического поля может быть равной до $2400/f$ [кГц] мкВ/м на 300 м, где f – частота. В диапазоне 90 кГц – 110 кГц ограничивается только паразитное излучение. В диапазоне 160 кГц – 190 кГц имеются различные назначения. Таким образом для систем RFID доступны диапазоны частот 9 – 90 кГц и 110 – 160 кГц. В диапазонах частот 9 – 90 кГц и 110 – 490 кГц измерения проводятся с использованием детектора среднего значения.

Таблица 4.3. Частоты от 9 – 160 кГц, доступные для общего использования согласно регламенту FCC.

Полоса частот	Напряженность электрического поля	~ Напряженность магнитного поля	Детектор
9 – 90 кГц	$2400/f$ [кГц] мкВ/м на 300 м	$198,8 - 60 \log(f[\text{кГц}])$ дБмкА/м на 10 м	Среднего значения
110 – 160 кГц	$2400/f$ [кГц] мкВ/м на 300 м	$198,8 - 60 \log(f[\text{кГц}])$ дБмкА/м на 10 м	Среднего значения

В Японии регламенты определяются законом MPT. Согласно параграфам 1 и 3 статьи 4 радиосредства с очень низким уровнем излучения могут использоваться без лицензий. В статье 6 регламентируются соответствующие уровни излучения [42]. Для частот ниже 322 МГц напряженность электрического поля на расстоянии 3 м должна быть не более 500 мкВ/м. На частоте приблизительно 16 МГц расстояние 3 м является точкой перехода от ближней зоны к дальней зоне. На частотах менее 16 МГц точка 3 м лежит в ближней зоне. Поэтому для частот менее 16 МГц предел 500 мкВ/м задан в точке перехода от ближней к дальней зоне. Для того чтобы рассчитать магнитное поле в некоторой точке пространства, в нашем случае 10 м, необходимо использовать уравнение (4.3) или (4.7) в зависимости от того, находится расстояние в ближней или дальней зоне. Так как для частот менее 4,8 МГц расстояние 10 м лежит в ближней зоне, необходимо использовать уравнение (4.7). Для частот от 4,8 МГц до 15,9 МГц расстояние 10 м лежит в дальней зоне, поэтому необходимо использовать уравнение (4.3). Следует заметить, что эти рассуждения основаны на сравнении, приведенном в регламенте [11] и не применимы в случае применения MPT и ARIB стандартов.

Таблица 4.4. Частоты, доступные для использования радиосредствами низкого уровня излучения, согласно МРТ регламенту.

Полоса частот	Напряженность электрического поля	~ Напряженность магнитного поля
9 кГц – 4,8 МГц	500мкВ/м на (с/2πf) м	229,2 – 60 log(f[кГц]) мкА/м на 10м
4,8 МГц – 15,9 МГц	500мкВ/м на (с/2πf) м	76 – 20 log(f[кГц]) мкА/м на 10м
15,9 МГц – 322 МГц	500мкВ/м на 3 м	–8 дБмкА/м на 10м

4.6.1.2. Диапазон 6,78 МГц

Диапазон 6,695 – 6,795 МГц является диапазоном ISM, разрешенным в странах, присоединившихся к RR S5.138 [36]. Также он используется для радиовещания и аэрокосмической связи, кроме устройств ISM. Регламенты в Европе этот диапазон относят к предпочтительным для систем RFID. Это, однако, не распространяется на США.

СЕРТ 70-03 в приложении 9 дает ограничения для индуктивных применений [44]. Предельная напряженность магнитного поля равна 42 дБмкА/м на 10 м. В пределах +/- 150 кГц от центральной частоты 6,78 МГц предельная напряженность магнитного поля равна 9 дБмкА/м на 10 м. Вне этих частот предельная напряженность магнитного поля снижается до -1 дБмкА/м на 10 м. Измерения производятся при помощи квазипикового детектора.

Таблица 4.5. Ширина спектра и напряженность поля в диапазоне 6,78 МГц согласно регламенту СЕРТ/ETSI.

Полоса частот	Напряженность магнитного поля	Детектор
6,78 МГц +/- 30 кГц	42 дБмкА/м на 10 м	Квазипиковый
6,78 МГц +/- 30 ... 150 кГц	9 дБмкА/м на 10 м	Квазипиковый
6,78 МГц +/- 150 кГц →	–1 дБмкА/м на 10 м	Квазипиковый

Согласно FCC 15.233 в диапазоне 1,705 – 10 МГц определено, что если ширина спектра сигнала по уровню 6 дБ превышает или равна 10% от центральной частоты, предельный уровень электрического поля устанавливается равным 100 мкВ/м на 30 м. Если ширина спектра сигнала по уровню 6 дБ меньше 10% от центральной частоты, предельный уровень электрического поля устанавливается равным 15 мкВ/м на 30 м или величине в мкВ/м, соответствующей ширине спектра устройства в килогерцах, поделенной на центральную частоту в мегагерцах. Для устройств ISM ширина спектра должна быть меньше 30 кГц. Предельный уровень напряженности электрического поля равен 15 мкВ/м на 30 м [46]. Это, однако, не является обязательным условием.

Таблица 4.6. Ширина спектра и напряженность поля в диапазоне 6,78 МГц согласно FCC регламенту.

Полоса частот	Напряженность электрического поля	~ Напряженность магнитного поля	Детектор
6,78 МГц +/- 30 кГц или	15 мкВ/м на 30 м	-18,5 дБмкА/м на 10м	Среднего значения
6,78 МГц +/- 678 кГц	100 мкВ/м на 30 м	-1,98 дБмкА/м на 10м	Среднего значения

В Японии приняты нормы, которые приблизительно эквивалентны данным, приведенным в табл. 4.4. [46].

4.6.1.3. Диапазон 13,56 МГц

Диапазон 13,553 – 13,567 МГц во всем мире является диапазоном ISM (RR S5 150). Кроме большого числа SRD и ISM устройств, этот диапазон используют трансконтинентальные службы. Для систем RFID регламенты во всем мире относят диапазон HF - 13,56 МГц к предпочтительным.

Ограничения для индуктивных применений даны в СЕРТ 70-03 приложение 9 [44]. Предельная напряженность магнитного поля равна 42 дБмкА/м на 10 м. В пределах +/- 150 кГц от центральной частоты 13,56 МГц предельная напряженность магнитного поля равна 9 дБмкА/м на 10 м. Вне этих частот предельная напряженность магнитного поля снижается до -3,5 дБмкА/м на 10 м. Измерения производятся при помощи квазипикового детектора.

Таблица 4.7. Ширина спектра и напряженность поля в диапазоне 13,56 МГц согласно регламентам СЕРТ/ETSI.

Полоса частот	Напряженность магнитного поля	Детектор
13,56 МГц +/- 7 кГц	42 дБмкА/м на 10м	Квазипиковый
13,56 МГц +/- 30-150 кГц	9 дБмкА/м на 10м	Квазипиковый
13,56 МГц +/- 150 кГц→	-3,5 дБмкА/м на 10м	Квазипиковый

Согласно FCC 15.233 в диапазоне ISM определен предел напряженности электрического поля, равный 10 000 мкВ/м на 30 метрах. Вне диапазона предельная напряженность электрического поля снижается до 30 мкВ/м на 30 м. Из-за требований радиоастрономии ниже 13,41 МГц допускается только паразитное излучение. Внутри диапазона измерения проводятся при помощи квазипикового детектора, а за его пределами – при помощи детектора среднего значения [46].

Таблица 4.8. Ширина спектра и напряженность поля в диапазоне 13,56 МГц согласно FCC регламенту.

Полоса частот	Напряженность электрического поля	~ Напряженность магнитного поля	Детектор
13,56 МГц +/- 7 кГц	10 000 мкВ/м на 30 м	38,0 дБмкА/м на 10 м	Квазипиковый
13,56 МГц +/- 7 кГц →	30 мкВ/м на 30 м	-12,4 дБмкА/м на 10 м	Ср. значения

В Японии системы RFID, работающие в этом диапазоне, могут занимать полосу, равную 7-кратной модуляционной скорости, при мощности на выходе антенны, равной 10 мВт и ее усилении равном или меньшем 30 дБи. Такое требование эквивалентно 10 Вт EIRP и напряженности магнитного поля 73,2 дБмкА/м на 10 м. При усилении антенны 2,15 дБи это также соответствует напряженности магнитного поля 45,4 дБмкА/м на 10 м [42]. Такие нормы должны быть подтверждены соответствующими документами, которые, однако, до настоящего времени не опубликованы.

4.6.1.4. Диапазон 27 МГц

Во всем мире диапазон 26,957 – 27,283 МГц принят в качестве ISM диапазона. Он используется большим числом устройств небольшой дальности и устройствами ISM, включая нагревательные и медицинские установки, пейджеры и устройства управления радиомоделями. В Северной Америке и Европе он лежит внутри общедоступного СВ (Citizens Band) радиодиапазона. Диапазон 27 МГц, хотя и имеет достаточно большую ширину, редко используется для RFID.

В пределах ISM диапазона СЕРТ в рекомендациях 70-03 допускает напряженность магнитного поля до 42 дБмкА/м на расстоянии 10 м при использовании квазипикового детектора или 10 мВт ERP выходной мощности (Приложение 9) [44].

Таблица 4.9. Ширина спектра и напряженность поля в диапазоне 27,12 МГц согласно СЕРТ/ETSI регламентам.

Полоса частот	Напряженность магнитного поля	Детектор
27,12 МГц +/- 163 кГц	42 дБмкА/м на 10м	Квазипиковый
27,12 МГц +/- 163 кГц →	- 3,5 дБмкА/м на 10м	Квазипиковый

В пределах ISM диапазона FCC (15.225) устанавливает максимальную напряженность электрического поля до 10 000 мкВ/м на 30 м. Вне ISM диапазона предельная напряженность электрического поля установлена 30 мкВ/м на 30 м (15.209). Измерения должны проводиться при помощи детектора среднего значения [46].

Таблица 4.10. Ширина спектра и напряженность поля в диапазоне 27,12 МГц согласно FCC регламенту.

Полоса частот	Напряженность электрического поля	~ Напряженность магнитного поля	Детектор
27,12 МГц +/- 163 кГц, или	10 000 мкВ/м на 30 м	18,0 дБмкА/м на 10 м	Среднего значения
27,12 МГц +/- 163 кГц →	30 мкВ/м на 30 м	-12,4 дБмкА/м на 10 м	Среднего значения

В Японии установлен общий предел 500 мкВ/м на 30 м (- 8 дБмкА на 10 м) [42].

4.6.1.5. Диапазон 40,68 МГц

Во всем мире диапазон 40,66 – 40,70 МГц также принят в качестве ISM диапазона. Он редко используется для RFID, так как характеристики распространения и регламентные ограничения в этом диапазоне значительно менее благоприятны, чем в диапазонах 13,56 МГц и 27,12 МГц.

CEPT/ETSI устанавливает максимальную выходную мощность, равную 10 мВт ERP в пределах ISM диапазона. Вне ISM диапазона предельный уровень падает до 250 нВт ERP [44,49].

Таблица 4.11. Ширина спектра и напряженность поля в диапазоне 40,68 МГц согласно регламентам CEPT/ETSI.

Полоса частот	Мощность (ERP)	~ Напряженность магнитного поля	Детектор
40,7 МГц +/- 20 кГц	10 мВт	45,4 дБмкА/м на 10 м	Квазипиковый
40,7 МГц +/- 20 кГц →	250 нВт	- 0,62 дБмкА/м на 10 м	Квазипиковый

Пределы FCC изменяются в зависимости от типа устройств. Для устройств практически непрерывного излучения предельная напряженность электрического поля установлена 1000 мкВ/м на 3 м при измерении квазипиковым детектором. Вне диапазона установлен предел 100 мкВ/м на 3 м (15.229) при измерении детектором среднего значения [46].

Периодически излучающие устройства могут излучать до 2250 мкВ/м на 3 м. Паразитное излучение не должно превышать 225 мкВ/м. Продолжительность излучения не должна превышать 5 с. Измерения могут проводиться или квазипиковым или детектором среднего значения.

Таблица 4.12. Ширина спектра и напряженность поля в диапазоне 40,68 МГц согласно регламенту FCC.

Полоса частот	Напряженность электрического поля	~ Напряженность магнитного поля	Детектор
40,7 МГц +/- 20 кГц	1000 мкВ/м на 3 м	- 1,98 дБмкА/м на 10 м	Среднего значения
40,7 МГц +/- 20 кГц →	100 мкВ/м на 3 м	- 22,0 дБмкА/м на 10 м	Ср. зн.

В Японии установлен общий предел излучения 500 мкВ/м на 30 м (– 8 дБмкА на 10 м).

4.6.2. Частоты RFID в дальней зоне

На UHF и микроволновых частотах предельные нормы интенсивности излучения обычно выражаются в единицах излучаемой мощности ERP или EIRP.

В диапазонах 900 МГц, 2,45 ГГц и 5,8 ГГц допускается работа устройств в режимах перестройки частоты FH и/или DS. Как правило администрации разрешают функционирование систем и с распределенным спектром, и в узкополосном режиме.

Необходимо отметить, что в настоящее время во всем мире в диапазонах 900 и 2450 МГц предлагается к внесению большое число изменений в регламенты. Мы рассмотрим как существующее положение, так и предлагаемые изменения.

Кроме того, с учетом определенной специфики, в Японии нормы задаются в виде пределов излучения для радиосредств малой мощности. Так для частот от 322 МГц до 10 ГГц предельная напряженность электрического поля равна 35 мкВ/м на 3 м [42].

Таблица 4.13. Нормы излучения для радиосредств малой мощности в диапазоне от 322 МГц до 10 ГГц согласно регламенту MPT.

Полоса частот	Напряженность электрического поля	Мощность (ERP)
322 МГц – 10 ГГц	35 мкВ/м на 3м	0,37 нВт

4.6.2.1. Диапазон 433 МГц

Диапазон 433,050 – 434,790 МГц не является диапазоном ISM. В соответствии RR S5.280 он разрешен в 10 европейских странах и в соответствии RR S5.138 – в Регионе 1 [36]. Диапазон 433 МГц лежит в пределах радиоловительского диапазона и широко используется в различных SRD и ISM устройствах.

В Европе CEPT/ETSI разрешает этот диапазон для неспецифических SRD устройств [44]. Мощность ERP не должна превышать 10 мВт, а скважность - 10%.

Таблица 4.14. Ширина спектра и пределы излучаемой мощности в диапазоне 433 МГц согласно регламенту СЕРТ/ETSI.

Полоса частот	Мощность (ERP)	Мощность (EIRP)	Скважность	Детектор
433,92 МГц +/- 870 кГц	10 мВт	16,4 мВт	< 10%	Квазипиковый
433,92 МГц +/- 870 кГц→	250 нВт	410 нВт	< 10%	Квазипиковый

В качестве ISM этот диапазон не разрешен ни в США, ни в Японии, хотя в Японии он регламентируется и используется в качестве диапазона для радиосредств малой мощности.

4.6.2.2. Диапазон 862 – 870 МГц

Диапазон 868 – 870 МГц не является диапазоном ISM, хотя он и разрешен СЕРТ для неспецифических устройств SRD. Прилежащие к этому диапазону частоты используются в беспроводной и сотовой телефонии, которые чувствительны к помехам. Ситуация в этом диапазоне сложна и прорабатываются различные предложения по дополнительному выделению для устройств SRD частот в области 862 – 868 МГц. Мы рассмотрим и текущее, и перспективное положение в этом диапазоне.

В настоящее время диапазон 868 – 870 МГц разбит на пять поддиапазонов: 868 – 868,6 МГц, 868,7 – 869,2 МГц, 869,3 – 869,4 МГц, 869,4 – 869,65 МГц и 869,7 – 870 МГц. Из всех поддиапазонов один – 869,4 – 869,65 МГц, допускает наибольший уровень мощности в 500 мВт ERP и, поэтому, наиболее приемлем для RFID. Полоса равна 250 кГц, а скважность не должна превышать 10%. Вне этого диапазона уровень ERP излучения должен быть не более 250 нВт [44].

Таблица 4.15. Ширина спектра и пределы излучаемой мощности на частоте 869,525 МГц согласно регламенту СЕРТ/ETSI в настоящее время.

Частота	Мощность (ERP)	~ Мощность (EIRP)	Скважность	Детектор
869,525 МГц +/- 125 кГц	500 мВт	820 мВт	< 10%	Квазипиковый
869,525 МГц +/- 125 кГц→	250 нВт	410 нВт	< 10%	Квазипиковый

Так как европейские нормы значительно ниже, чем в США, рассматривается предложение для неспецифических устройств SRD снизить требования в диапазоне 865 – 868 МГц. Предлагается, чтобы устройства RFID в этом диапазоне функционировали либо на фиксированных частотах, либо с дискретной перестройкой частоты на 15 каналах по 200 кГц. Для того чтобы предотвратить помехи в соседних диапазонах, предлагается снижение предела мощности на краях диапазона. Эти предложения представлены в табл. 4.16.

Таблица 4.16. Предлагаемая ширина спектра и пределы излучаемой мощности в диапазоне 865 – 868 МГц согласно регламенту CEPT/ETSI.

Частота	Мощность (ERP)	~ Мощность (EIRP)	Детектор
865 – 865,6 МГц	100мВт	164 мВт	Квазипиковый
865,6 – 867,6 МГц	2Вт	3,28 Вт	Квазипиковый
867,6 – 868 МГц	500 мВт	820 мВт	Квазипиковый
865 – 868 МГц→	250 нВт	410 нВт	Квазипиковый

4.6.2.3. Диапазон 915 МГц

В Регионе 2 частоты 902 - 928 МГц назначены в качестве диапазона ISM. Этого нет ни в Европе, ни в Японии. Однако Канада, Австралия, Новая Зеландия и большинство стран Латинской Америки разрешают работать или в этом, или в близких диапазонах. Диапазон 915 МГц широко используется большим числом устройств, включая сенсоры движения и беспроводные телефоны.

В США существует три направления использования этого диапазона. Первое направление ограничивается применением в сенсорах охранных устройств, второе – для работы радиоустройств с распределенным спектром SS в режимах прямой последовательности DS и скачкообразной перестройки частоты FH, и третье – для узкополосных применений. Мы рассмотрим только режим FH и узкополосное применение, хотя, как отмечалось в главе 3, используется и режим DS.

В диапазоне 915 МГц общие требования для систем с FH заданы FCC 15.24. Эти нормы применяются также в диапазонах 2,45 и 5,8 ГГц. Частоты каналов должны быть разнесены более чем на 25 кГц или на ширину полосы 20 дБ. Перескок частоты должен быть табличным, псевдослучайным, а использование каждой частоты должно быть в среднем одинаковым. Входная полоса приемника должна быть согласована с шириной спектра излучения и должна перестраиваться в соответствии с частотой передатчика. При отстройке на 100 кГц мощность должна быть на 20 дБ ниже мощности на основной частоте. Спектральные компоненты вне диапазона должны соответствовать FCC 15.209. В системах FH отсутствует необходимость перескока частоты во всем диапазоне в течение каждого сеанса излучения, однако, если сеанс продолжителен, перескоки должны быть. Для того чтобы определить и избежать появления других пользователей на частоте канала, возможен предварительный анализ занятости частоты, однако других специальных мер для избежания возникновения коллизий не требуется [46].

В диапазоне 902 – 928 МГц каналы могут занимать или до 250 кГц полосы 20 дБ, или от 250 до 500 кГц. Системы, которые используют полосу до 250 кГц, должны иметь, по крайней мере, 50 перескоков частоты и среднее время работы на канале 0,4 с в течение 20 с. Системы,

использующие полосу от 250 кГц до 500 кГц, должны иметь, по крайней мере, 25 перескоков частоты и среднее время работы на канале 0,4 с в течение 10 с.

Таблица 4.17. Ширина полосы и пределы мощности для систем FH в диапазоне частот 902–928 МГц согласно регламенту FCC.

Частота	Диапазон	Каналы	Мощность (EIRP)
902 – 928 МГц	< 250 кГц	≥ 50	4 Вт
902 – 928 МГц	≥ 250 кГц, ≤ 500 кГц	≥ 25	1 Вт

Системы, имеющие, по крайней мере, 50 перескоков частоты, могут работать с пиковой мощностью на входе антенны до 1 Вт. Системы с числом перескоков от 25 до 50 могут работать с пиковой мощностью на входе антенны до 0,25 Вт. Если используются антенны с усилением более 6 дБи, соответственно должна быть снижена пиковая мощность на входе антенны. При этом мощность EIRP должна сохраняться.

Согласно FCC 15.249 в диапазоне 902 – 928 МГц также могут работать устройства без перескока частоты. Напряженность поля на основной частоте не должна превышать 50 В/м, а на гармониках – 500 мкВ/м на 3 м. Вне диапазона излучение не должно превышать 500 мкВ/м на 3 м.

Таблица 4.18. Пределы мощности для узкополосного использования в диапазоне 902 – 928 МГц согласно регламенту FCC.

Частота	Напряженность электрического поля	~ Мощность (EIRP)	Детектор
902 – 928 МГц (основная частота)	50 мВ/м на 3м	750 мкВт	Квазипиковый
902 – 928 МГц (гармоники)	500 мкВ/м на 3м	75 нВт	Квазипиковый
902 – 928 МГц →	500 мкВ/м на 3м	75 нВт	Среднего значения

4.6.2.4. Диапазон 2,45 ГГц

Во всем мире этот диапазон принят в качестве диапазона ISM и поэтому широко используется большим числом устройств. Основными источниками помех здесь являются СВЧ устройства, другие типы ISM устройств, а также беспроводная связь [51].

В Европе для систем RFID, работающих в этом диапазоне, параметры оговариваются специально [50]. В соответствии с CEPT 70-03 мощность неспецифических устройств SRD в диапазоне 2,4 – 2,4835 ГГц не может превышать 10 мВт EIRP [44], однако в дополнении 11 указано, что для применений RFID этот диапазон может использоваться с 2001 года. В настоящее время проект ETSI EN 300440-1 дает специальную маску для систем RFID, работающих в этом диапазоне. Так в диапазоне 2,446 – 2,454 ГГц мощность EIRP не должна превышать 27 дБм (≈ 500 мВт). С таким

уровнем мощности аппаратура в режимах FH и CW может использоваться как внутри, так и вне помещений. В этом же диапазоне для систем с перескоком частоты внутри помещений предел мощности EIRP равен 36 дБм (≈ 4 Вт). Также определено, что при более высоком уровне мощности, ширина ДНА не должна превышать 90° в горизонтальной плоскости и иметь уровень боковых лепестков не более -15 дБ. Перескок частоты должен иметь не менее 20 каналов со средним временем работы на канале не более 0,4 с [50].

Таблица 4.19. Диапазоны и пределы мощности для систем с перескоком частоты в диапазоне 2,45 ГГц согласно регламенту CEPT/ETSI.

Частота	Мощность (EIRP)	Скважность	Детектор
2,45 ГГц +/-4 МГц	4 Вт*	$\leq 15\%$ за 200 мс	Среднего значения
2,45 ГГц +/-4 МГц	500 мВт**	$\leq 100\%$	Среднего значения
2,45 ГГц +/-4 МГц →	1,64 мкВт		Среднего значения

* Только для систем FH внутри помещений.

** Для систем FH или CW внутри и снаружи помещений.

В США системы RFID в этом диапазоне могут работать как с перескоком частоты, так и без него. Общие требования к системам с перескоком частоты в этом диапазоне совпадают с требованиями в диапазоне 902 – 928 МГц. Однако ширина диапазона, число каналов и другие детали отличаются. В этом диапазоне системы могут работать по крайней мере на 75 каналах или, в особых случаях по крайней мере на 15 каналах. При работе на 75 каналах полоса на уровне 20 дБ не должна превышать 1 МГц, а среднее время работы на одной частоте не должно превышать 0,4 с за период 30 с. Если система использует по крайней мере 15 не перекрывающихся каналов с общим перекрытием 75 МГц, полоса на уровне 20 дБ может превышать 1 МГц. В этом режиме среднее время работы на одной частоте должно быть равным 0,4 с за время, которое занимает перестройка частоты во всем диапазоне. При работе на 75 каналах мощность на выходе антенны не должна превышать 1 Вт. При работе с большим, чем 75, числом каналов, она не должна превышать 0,125 Вт. Если усиление антенны превышает 6 дБи, мощность на выходе антенны должна уменьшаться на соответствующую величину. Требования к фиксированной связи «точка – точка» немного отличны [46].

Таблица 4.20. Диапазоны и пределы мощности для систем с перескоком частоты в диапазоне 2,45 ГГц согласно регламенту FCC.

Частота	Число каналов	Ширина полосы (20 дБ)	Мощность (EIRP)	Детектор
2,4 – 2,4835 ГГц	≤ 75	1 МГц	4 Вт	Среднего значения
2,4 – 2,4835 ГГц	≤ 15	> 1 МГц (перекрытие 75 МГц)	500 мВт	Среднего значения
2,4 – 2,4835 ГГц →			75 нВт	Среднего значения

Согласно FCC 15.249 системы в этом диапазоне могут работать и без перестройки частоты. Основным ограничением является неперевышение напряженности основного излучения 50 мВ/м на 3 м и неперевышение напряженности поля на гармониках 500 мкВ/м на 3 м. Измерения должны проводиться с использованием детектора среднего значения при условии, что пиковая напряженность поля не должна превышать установленные пределы более чем на 20 дБ.

Таблица 4.21. Диапазоны и пределы мощности для узкополосного использования в диапазоне 2,45 ГГц согласно регламенту FCC.

Частота	Напряженность электрического поля	~ Мощность (EIRP)*	Детектор
2,4 – 2,4835 ГГц (основное излучение)	50 мВ/м на 3 м	750 мкВ	Среднего значения
2,4 – 2,4835 ГГц (гармоники)	500 мкВ/м на 3 м	75 нВ	Среднего значения
2,4 – 2,4835 ГГц →	500 мкВ/м на 3 м	75 нВ	Среднего значения

Стандарты для радиоэлектронных устройств (RCR STD), введенные Ассоциацией радиопромышленности и бизнеса ARIB, в документах RCR STD-1 и STD-29 устанавливают пределы для систем RFID диапазона 2,45 ГГц. STD-1 устанавливает пределы для автомобильных идентификационных систем, а STD-29 – для радиостанций малой мощности. STD-29 устанавливает пределы для узкополосных устройств с выходной мощностью антенны 10 мВт и усилением 20 дБи. Предполагается, что STD-29 установит нормы для систем с перестройкой частоты. Пределы могут быть эквивалентными существующим пределам для беспроводных систем локальной связи – мощность 10 мВт и усиление антенны 2,14 дБи [42].

Таблица 4.22. Диапазоны и пределы мощности в диапазоне 2,45 ГГц согласно регламенту МРТ.

Частота	Мощность/ усиление	Мощность (EIRP)
2,4 – 2,4835 ГГц (узкополосный режим)	10 мВт/20 дБи	1 Вт EIRP
2,4–2,4835 ГГц (с перестройкой частоты)	270 мВ/2,14 дБи	0,44 Вт

4.6.2.5. Диапазон 5,78 ГГц.

Этот диапазон во всем мире принят в качестве диапазона ISM. В последнее время он загружен локальными беспроводными сетями, однако не так плотно, как диапазон 2,45 ГГц.

В Европе мощность EIRP неспецифических SRD не должна превышать 25 мВт. Специальных указаний по приложениям RFID в этом диапазоне нет. Паразитные излучения свыше 1 МГц должны быть менее 1 мкВт [50].

Таблица 4.23. Диапазоны и пределы мощности для неспецифических SRD в диапазоне 5,78 ГГц согласно регламенту CEPT/ETSI.

Частота	Мощность (EIRP)	Детектор
5,725 – 5,875 ГГц	25 мВт	Среднего значения
5,725 – 5,875 ГГц →	1 мкВт (или ERP)	Среднего значения

В США, как и в диапазонах 915 и 2450 МГц, устройства могут работать с распределенным спектром или в обычном узкополосном режиме. В диапазоне 2450 МГц к устройствам с перестройкой частоты применяются такие же требования, как и в диапазоне 915 МГц. Системы с перестройкой частоты в этом диапазоне могут работать по крайней мере на 75 каналах; при этом полоса на уровне 20 дБ не должна превышать 1 МГц, а среднее время работы на одной частоте не должно превышать 0,4 с за период 30 с. Мощность на выходе антенны не должна превышать 1 Вт, а если усиление антенны превышает 6 дБи, мощность должна уменьшаться на величину превышения. Требования к фиксированной связи «точка – точка» в диапазонах 915 и 2450 МГц различны [46].

Таблица 4.24. Диапазоны и пределы мощности для режима с перестройкой частоты в диапазоне 5,78 ГГц согласно регламенту CEPT/ETSI.

Частота	Число каналов	Полоса частот (20 дБ)	Мощность (EIRP)	Детектор
5,725 – 5,875 ГГц	≥75	1МГц	4Вт	Среднего значения
5,725 – 5,875 ГГц →			75 нВт	Среднего значения

В обычном узкополосном режиме согласно требованиям FCC 15.249, напряженность основного электрического поля не должна превышать 50 мВ/м, а напряженность на гармониках – 500 мкВ/м на 3 м. Эти нормы проверяются при помощи детектора среднего значения, а при использовании пикового детектора указанные нормы не должны превышать более чем на 20 дБ.

Таблица 4.25. Диапазоны и пределы мощности для узкополосного режима в диапазоне 5,78 ГГц согласно регламенту FCC.

Частота	Напряженность электрического поля	~ Мощность (EIRP)*	Детектор
5,725 – 5,875 ГГц (основное излучение)	50 мВ/м на 3 м	750 мкВ	Среднего значения
5,725 – 5,875 ГГц (гармоники)	500 мкВ/м на 3 м	75 нВ	Среднего значения
5,725 – 5,875 ГГц→	500 мкВ/м на 3 м	75 нВ	Среднего значения

В Японии напряженность излучения в диапазоне 5,78 ГГц не должна превышать 35 мкВ/м на расстоянии 3 м.

4.7. Резюме

Из рассмотрения возможных частотных диапазонов, предельных напряженностей поля и мощностей излучения в Европе, США и Японии можно сделать вывод, что для систем RFID наиболее подходят диапазоны < 135 кГц, 13,56 кГц и 2,45 ГГц. В Европе и США также приемлемы диапазоны 125 кГц и 900 МГц. Другие диапазоны могут использоваться только отдельными администрациями, например, диапазон 6,78 МГц удобен для использования в Европе, но не в США, и наоборот - диапазон 5,78 ГГц удобен для использования в США, но не в Европе.

При выборе диапазона частот должны учитываться не только доступность, предельная напряженность поля и мощность излучения, но и многие другие факторы. При этом необходимо учитывать не только распространение и поведение электромагнитных волн в окружающей среде, как это ранее рассматривалось в главе 2, но и вопросы проектирования и производства, а также конкретные пользовательские требования к аппаратуре RFID. В главе 5 мы рассмотрим перечисленные вопросы.

Глава 5

Аппаратурное исполнение

5.1. Введение

В главе 2 мы рассмотрели свойства электромагнитных полей и волн, а в главе 3 – теоретические вопросы, относящиеся к тому, как при помощи модуляции электромагнитных полей метки взаимодействуют со считывателем. В главе 4 мы увидели, как электромагнитные поля и их модуляция ограничиваются административными нормами - регламентами. Теперь мы рассмотрим, как технически реализуются основные функции метки при наличии взаимно накладывающихся ограничений.

Аппаратная реализация меток исключительно важна с точки зрения их характеристик и стоимости. Хотя оба эти фактора важны для реализации преимуществ технологии RFID, в первую очередь важна стоимость. Пассивные метки, так же, как и обычные товары повседневного спроса, которые мы покупаем, имеют конечный срок службы. Они будут производиться и утилизироваться также часто, как и покупаемые товары. При гигантском объеме производства и использования меток, минимальное повышение стоимости одной метки приведет для потребителей к огромному увеличению их общей стоимости. Характеристики метки определяются не только собственно ею, но и считывателем и связанной с ним инфраструктурой. С другой стороны, стоимость определяется только самой меткой – ее проектированием, изготовлением и объемом производства. По этой причине нам необходимо понять, как физические компоненты метки обеспечивают ее функционирование, и как это влияет на ее стоимость.

Метка состоит из антенны, соединенной линией передачи с электронной схемой. Антенна является преобразователем электромагнитного поля в электрическую энергию и наоборот. Линия передачи передает эту энергию в электронную схему и наоборот. Электронная схема преобразует, запасает и использует энергию, и, наконец, направляет ее обратно в линию передачи и антенну, вызывая управляемый отклик.

Прежде всего мы кратко рассмотрим антенны для меток. Затем мы рассмотрим электронные схемы, входящие в состав метки, включая их компоненты и энергопотребление. Основное внимание в этой главе будет сконцентрировано на стоимости и характеристиках, которые непосредственно связаны с проектированием и параметрами меток. Процесс производства, несомненно, очень важен и заслуживает подробного анализа. Однако этот вопрос мы затронем очень коротко.

5.2. Антенны

Существует два основных типа меток RFID. Одни предназначены для работы в ближней зоне, другие же работают в дальней зоне электромагнитного поля. Так как взаимодействие с электромагнитными полями и волнами у этих двух типов меток существенно отличается, то и антенны различны. Сначала мы рассмотрим антенны ближней зоны, а затем – дальней зоны поля.

5.2.1. Ближняя зона

Системы RFID в ближней зоне связаны либо индуктивным, либо емкостным способом. Индуктивно связанные системы взаимодействуют при помощи магнитного, а емкостные – электрического поля. Как мы видели, предпочтительным выбором для индуктивно связанных систем является рамочная антенна, так как излучение магнитного поля у нее преобладает над излучением электрического поля. Для систем с емкостной связью предпочтительными являются другие антенны – дипольные или электродные. Более подробно мы остановимся на индуктивно связанных системах с рамочными антеннами.

На характеристики и стоимость рамочных антенн влияет несколько факторов. С магнитным полем антенны непосредственно связаны диаметр и количество витков рамки – они определяют индуцированное напряжение и ток и, в результате, дальность действия системы. Большая площадь и большее количество витков приводят к большему расходу материалов и повышению стоимости антенны. Кроме того свойства материалов из-за их влияния на импеданс и добротность, влияют на характеристики системы и, в частности, на дальность действия. Различия в стоимости материалов также влияют на стоимость метки в целом. Кратко рассмотрим влияние размера и материалов, из которых изготовлена рамка, на основные технические характеристики метки и ее стоимость.

Анализируя уравнение (2.29), можно увидеть прямую зависимость индуцируемого напряжения от площади, которую охватывает рамка, числа витков и частоты:

$$V_{1 \rightarrow 2} = j\omega N_2 \mu H_1 A_2 \cos \psi .$$

Все прочие параметры, включая пространственную ориентацию, будем считать максимизированными, а проницаемость – соответствующей свободному пространству. Проницаемость μ можно повысить

использованием ферритового сердечника, при этом одновременно увеличивается направленность антенны.

Особенно важным является соотношение между частотой ω , числом витков N и площадью A , которую охватывает петля. Для максимизации индуцируемого напряжения рамка метки должна быть согласована с площадью рамки считывателя. Однако из-за существующих ограничений по размерам и стоимости материала, размеры рамки метки стремятся сделать минимальными. Следует отметить, что метки, работающие на низших частотах, на 125 кГц например, требуют при тех же характеристиках почти в 110 раз больше витков, чем на 13,56 МГц.

Число витков рамочной антенны также влияет на индуктивность и потери рамки. В качестве примера мы можем рассмотреть круглую рамку (прямоугольная рамка дает аналогичный результат). Индуктивность круглой рамки описывается выражением (2.31):

$$L_2 = \mu b N^2 [\ln(8b/a) - 1,75].$$

Заметим также, что индуктивность возрастает с радиусом рамки или ее площадью. Сопротивление круглой рамки описывается выражением (2.33):

$$R_{L2} = N^2 (b/a) (\omega \mu / 2\sigma)^{1/2},$$

где σ – проводимость материала.

Таким образом, влияние площади рамки и числа витков на стоимость в основном определяется видом материала.

Рамочные антенны, как правило, изготавливаются в виде свернутого провода или печатной проводящей пасты. Вид рамки непосредственно определяет и ее характеристики и стоимость. От вида материала зависит индуктивность и сопротивление рамочной антенны. Для максимизации добротности Q резистивные потери должны быть минимальными, а индуктивность – максимальной. Это обеспечивает максимальное напряжение V , хотя большая добротность не всегда желательна из-за проявления эффекта расстройки контура окружающими предметами.

Когда необходимо большое число витков, обычно используется медный провод. Удельное сопротивление медного провода приблизительно равно $0,017 \text{ Ом}\cdot\text{мм}^2/\text{м}$. С уменьшением диаметра провода резистивные потери возрастают, а добротность Q – падает.

Могут использоваться печатные проводящие краски и пасты, однако их применение ограничено рамками с малым числом витков. Возможно использование различных токопроводящих красок, но наиболее часто используются серебросодержащие краски благодаря их высокой

проводимости при приемлемой стоимости. Также как в медном проводе, потери печатных антенн растут с уменьшением ширины линии. Поверхностное сопротивление серебряных проводящих паст при толщине 1 микрон варьируется от 0,01 до 0,05 Ом на квадрат. Сопротивление углесодержащих паст колеблется в пределах между 10 и 50 Ом на квадрат при той же толщине [52]. Это сопротивление зависит от температуры полимеризации пасты и других факторов.

Важными являются размеры подстроечного конденсатора для настройки антенного контура в резонанс. На низких частотах необходима большая емкость. При этом, соответственно, площадь конденсатора в интегрированной электронной схеме также оказывается достаточно большой, что приводит к существенному повышению стоимости метки.

Необходимо учитывать и импедансный элемент, используемый для модуляции нагрузки. Это может быть конденсатор или резистор, или их комбинация. Для снижения мощности потребления и размеров компонентов этот импеданс должен быть минимизирован при условии обеспечения достаточного уровня модулированного сигнала, что необходимо для его обнаружения считывателем.

5.2.2. Дальняя зона

В дальней зоне системы RFID взаимодействуют при помощи электромагнитных волн. Как мы видели, высокая эффективность антенны достигается в том случае, когда по крайней мере один из ее размеров порядка половины длины волны. При таких размерах антенна является резонатором и имеет реальный входной импеданс. Это обеспечивает повышение эффективности излучения и приема, облегчение согласования и передачи мощности от/к электронной схеме, которая подсоединена к антенне.

Следует заметить, что использование полуволновой антенны в системах, которые функционируют в дальней зоне, не является обязательным. Могут использоваться антенны меньших размеров, однако при этом реактивная часть импеданса увеличится, а активная – уменьшится. Это может вызвать сложности в согласовании антенны с нагрузкой. В результате уменьшится эффективность передачи мощности сигнала, что, в свою очередь, уменьшит доступную мощность метки и снизит дальность действия системы в целом.

Существует несколько типов резонансных антенн. Обычно используются фольгированные полуволновые диполи и микрополосковые (патчевые) антенны различных конфигураций. Дипольные антенны бывают проволочного типа или печатного типа – на подложке. Зачастую патчевые антенны печатаются на диэлектрике, в виде печатной схемы на материале с

проводящей поверхностью снизу. Кратко рассмотрим некоторые характеристики и стоимость упомянутых антенн.

5.2.2.1. Дипольные антенны

Обычно ввиду производственных особенностей проволочные диполи стоят дороже, чем печатные антенны, поэтому в применениях, требующих больших количеств, печатные антенны предпочтительнее. Если в антенне наиболее важны омические потери, ширина диполя должна быть достаточно большой. При этом следует использовать материалы высокого качества с малым сопротивлением.

Из-за необходимости обеспечения высокой эффективности за счет резонанса полуволновые диполи могут иметь относительно небольшую полосу пропускания. Полоса пропускания диполя может быть увеличена за счет увеличения ширины или эффективного диаметра вибраторов [13].

5.2.2.2. Микрополосковые (патчевые) антенны

Микрополосковые или патчевые антенны являются планарными антеннами, которые, как правило, наносятся печатным образом на диэлектрическое основание. Они состоят из трех слоев: сверху – проводящая пластина (патч), снизу – проводящее металлическое основание и диэлектрический слой между ними. Пластина, обычно, по размерам порядка нескольких долей длины волны, а основание имеет несколько большие размеры. В качестве диэлектрика используются широко распространенные материалы, применяемые для высокочастотных печатных плат, например, PTFE или другие. Диаграмма излучения микрополосковых антенн в основном формируется областью поверхности над основанием [53].

Пластина может иметь различные формы. Обычно используется круглая, эллиптическая или прямоугольная форма. Различные формы пластин используются для того, чтобы обеспечить вариацию различных параметров, включая входной импеданс, эффективность излучения и поляризацию. Подвод энергии к пластине осуществляется тремя способами: при помощи коротких линий передачи, непосредственно электромагнитной связью или апертурным способом. Каждый способ запитки имеет свои особенности, но, как и в случае формы пластины, основными параметрами, на которые влияет запитка, остаются входной импеданс, эффективность излучения и поляризация.

Обычно в патчевых антеннах ширина полосы B равна единицам процентов. Для прямоугольной конфигурации излучающей пластины над диэлектриком относительная полоса пропускания равна [13]:

$$B = 3,77 [(\varepsilon_r - 1) / \varepsilon_r^2] (W/L) (t/\lambda), \quad (5.1)$$

где ε_r – относительная диэлектрическая проницаемость, t – толщина диэлектрика, а W и L – габариты пластины. Очевидно, что увеличение толщины диэлектрика расширяет полосу пропускания антенны, а увеличение диэлектрической проницаемости уменьшает ее.

Выбор диэлектрика важен не только с точки зрения ширины полосы пропускания, но также важен для других характеристик антенны. Так для обеспечения ее хороших эксплуатационных характеристик диэлектрики должны иметь минимальные потери, высокую температурную стабильность и стабильность геометрических размеров, а также малые неоднородности.

Микрополосковые антенны могут иметь большую стоимость, чем стоимость печатных дипольных антенн, так как необходимы металлическое основание и диэлектрик. Печатные диполи также расположены на основании, но свойства их подложки значительно меньше влияют на характеристики антенны, чем в случае основания патчевой антенны.

В общем случае параметры патчевых и печатных дипольных антенн сравнимы, а их использование во многих случаях определяется конкретным применением, диапазоном частот, стоимостью материалов и производственными возможностями.

5.3. Интегральные схемы

По причинам, изложенным во введении, мы предполагаем, что в метке используется интегрированная электронная (интегральная) схема – чип, причем единственная интегральная схема. Часто может быть необходимым подключение к чипу дополнительных внешних дискретных элементов. В приложениях, предполагающих очень малые габариты и объем, по причинам снижения стоимости и производственных затрат, метка должна содержать только одну интегральную схему.

Целью этого раздела является не проектирование интегральной схемы, а описание различных компонентов, необходимых для ее функционирования, и оценка их влияния на характеристики и стоимость. Наиболее важными характеристиками чипа являются – потребляемая мощность, размеры и стоимость. Из-за того, что доступная метке мощность извлекается из электромагнитного поля, энергетический ресурс ограничен. Потребляемую мощность необходимо уменьшать, но таким образом, чтобы это не приводило к значительному ухудшению параметров системы, особенно дальности. Размеры компонентов непосредственно влияют на стоимость. В общем случае, чем больше используется кремния, тем дороже метка. Для того чтобы минимизировать стоимость, необходимо уменьшать

число компонентов и их размеры. Следует заметить, однако, что максимальное использование элементов чипа может привести к повышению числа переключений этих элементов, что напрямую приводит к повышению потребляемой мощности. Эти компромиссные вопросы мы рассмотрим в следующем разделе.

5.3.1. Общая схема построения

До описания различных компонентов, входящих в интегральную схему, мы коротко рассмотрим общую схему построения пассивной метки RFID. Не зависимо от того, работают ли метки в ближней или в дальней зоне поля, схемы их построения похожи.

Метка, являющаяся простым носителем информации, выполняет несколько основных функций. Она принимает и выпрямляет входной высокочастотный сигнал, выделяя энергию и информацию. Она хранит и обеспечивает энергопитание. Для запуска цифровых схем метка из полученной информации выделяет синхронизирующий сигнал. При помощи цифровых схем метка обрабатывает информацию и обеспечивает модуляцию входного сигнала, как это описано в главах 2 и 3.

Входная радиочастотная цепь является двунаправленным интерфейсным элементом, который включен между антенной и остальной электронной частью метки. Во входной цепи осуществляется выделение энергии и данных, а также их передача к цепям источника питания, смещения, выделения синхросигнала и обработки данных [54, 55]. Во входной цепи, кроме того, реализуется защита от возможной перегрузки, вызванной повышением уровня сигнала при изменении расстояния до считывателя. При помощи антенны с использованием модуляции по командам процессора информация передается обратно к считывателю.

Цепи источника питания, смещения, выделения синхросигнала и обработки данных, а также память метки реализуются при помощи аналоговых и цифровых компонентов. В случае только читаемых меток это, обычно, только читаемая память (ROM) или записываемая один раз и читаемая много раз память (WORM).

5.3.2. Компоненты

Компоненты, необходимые для обеспечения функций метки, включают аналоговые и цифровые цепи из пассивных и активных элементов, а также память. Эти компоненты могут быть включены в единую интегральную схему. Так как в процессе изготовления цифровых элементов используется специфическая технология, дополнительное

включение аналоговых элементов может быть затруднительным. При этом сложностью является и то, что аналоговые цепи подвержены изменению параметров от температуры и времени [56]. Кроме того, как мы увидим, важные пассивные элементы могут занимать достаточно большую площадь и иметь разброс параметров.

После краткого описания физики полупроводников, мы рассмотрим пассивные компоненты. Начнем с транзисторов, затем перейдем к пассивным компонентам, диодам и другим важным элементам, включая генераторы, схемы ФАПЧ, цифровые схемы и память. В конечном итоге мы рассмотрим общее энергопотребление чипа. Как и отмечалось ранее, мы не будем касаться производственных вопросов.

5.3.2.1. Полупроводники и транзисторы

Полупроводниками называются материалы, которые не являются ни хорошими проводниками, ни хорошими изоляторами. Они изготавливаются введением контролируемого количества атомов примеси в технологическом процессе, называемом легированием. Легированием области материала другим материалом, имеющим больше валентных электронов, получают область "n - типа". Наоборот, легированием области материала другим материалом, имеющим меньше валентных электронов, получают область "p - типа". Приложенное напряжения вызывает движение избыточных электронов. По многим различным причинам для изготовления полупроводников наиболее широко используется кремний. Он широко распространен и его электрические свойства относительно слабо зависят от температуры из-за большой ширины запрещенной зоны. Для обеспечения быстрого действия приборов иногда применяется арсенид галлия – он позволяет вдвое повысить быстродействие, но, однако, дороже обычного кремния и менее распространен в природе [57, 58].

Полупроводниковые устройства – транзисторы, были созданы в конце 40-х годов. Сначала были разработаны биполярные, а затем – металлооксидные (MOS) устройства. Технология MOS позволила уменьшить размеры и снизить энергопотребление. Первыми появились приборы p-типа (PMOS), а затем – n-типа (NMOS). Для дальнейшего снижения потребляемой мощности были разработаны комплиментарные (CMOS) приборы, однако за счет увеличения размеров. Приборы biCMOS объединили высокое быстродействие биполярной технологии с низким энергопотреблением технологии CMOS. В настоящее время на рынке полупроводниковых технологий преобладает CMOS технология и в обозримом будущем предполагается, что она сохранит существенное ценовое преимущество.

Упрощенно MOS устройство можно представить в виде параллельного пластинчатого конденсатора. Простым примером является NMOS прибор. Затвор, обычно являющийся сильно легированным полисиликоном, является одной пластиной конденсатора, которая отделена диэлектриком от полупроводника – другой пластины. Затвор расположен между двумя сильно легированными областями n-типа, одна из которых является истоком, а другая – стоком. Все остальные области в таком полупроводниковом приборе являются областями p-типа.

Когда к затвору не приложено напряжение, устройство не проводит ток. Когда к затвору приложено напряжение, создается электрическое поле, которое индуцирует в полупроводнике заряды противоположной полярности. Если приложить достаточное напряжение, заряды отрываются от поверхности полупроводника, а электроны истока и стока создают инверсный слой. Под воздействием электрического поля заряды движутся, а движение электронов создает ток. Вариация геометрических размеров и напряжений позволяет создать большое число разновидностей MOS приборов.

MOS устройства обладают как статическими, так и динамическими эффектами. Динамические эффекты определяются емкостями различных элементов устройства. Исток и сток образуют емкость из обратного смещенного перехода и подложки. Кроме того, между затвором и истоком, между затвором и стоком, а также между затвором и каналом и между каналом и объемом материала тоже образуются параллельные пластинчатые емкости. Комплементарные металлооксидные (CMOS) приборы являются комбинацией PMOS и NMOS приборов. При этом обеспечивается уменьшение потребляемой мощности. Транзисторы являются основным компонентом интегральных микросхем, так как они могут выступать как в роли пассивных компонентов, таких как конденсаторы и резисторы, так и в качестве активных компонентов.

5.3.2.2. Пассивные компоненты

Пассивные компоненты особенно важны в радиочастотных входных цепях и цепях питания меток RFID. Коротко рассмотрим использование этих компонентов в интегральных схемах.

5.3.2.2.1. Резисторы

Существует несколько способов создания резисторов при помощи CMOS технологии. В соединительных цепях часто используются такие материалы, как поликремний, металлы или другие более специализированные материалы. По сравнению с металлами поликремний обладает меньшей проводимостью – он имеет сопротивление

приблизительно 5 – 10 Ом на квадрат с разбросом около 35%. Он также имеет низкую удельную по площади паразитную емкость, низкий средний температурный коэффициент и коэффициент напряжения. Металлы характеризуются сопротивлением порядка 10 мОм на квадрат [60].

Когда требуется высокая резистивность, могут использоваться wells - элементы. Они имеют порядка 1 – 10 кОм на квадрат, однако обладают высокой паразитной емкостью по отношению к подложке. Кроме того, у них высокий начальный разброс, достигающий 50 – 80 %, большой температурный коэффициент и большой коэффициент напряжения.

В качестве малогабаритных резисторов также могут использоваться MOS транзисторы. Они, однако, имеют большой разброс и большой температурный коэффициент из-за температурной зависимости от подвижности электронов. Другим способом, который обеспечивает близкие к поликремнию параметры – проводимость и температурный коэффициент, является использование диффузии области сток-исток р-типа или n-типа. При этом, однако, существенным негативным фактором является большая паразитная емкость и повышенный коэффициент напряжения (нелинейность) [60].

5.3.2.2.2. Конденсаторы

Конденсаторы могут быть образованы различными способами. Общепринятым способом является использование двух параллельных пластин с промежуточным материалом. Так как толщина промежуточного диэлектрика достаточно велика, удельная по площади емкость обычно мала (5×10^{-5} пФ/мкм²). При этом вредное влияние оказывают паразитные емкости слоев, расположенных непосредственно над и под пластинами; они могут составлять 10 – 30% от основной емкости.

Для того чтобы создать емкости за счет боковых линий вдоль параллельных плоскостей, можно использовать промежуточные материалы. Такой прием может обеспечить 10-кратное увеличение емкости [60].

Так как емкость затвора является общей для транзистора, могут также использоваться MOS устройства. В зависимости от толщины диэлектрика могут быть достигнуты удельные емкости 1–5 фФ/мкм².

С использованием емкости р-n перехода могут быть реализованы другие методы. При этом из-за того, что емкость перехода зависит от его смещения, реализуется подстраиваемый конденсатор [60].

5.3.2.2.3. Индуктивности

При помощи CMOS устройств индуктивности с низким уровнем шума, искажениями и потребляемой мощностью, реализуются трудно.

Наиболее широко используются плоские спирали, однако даже для малых индуктивностей они требуют большой площади, а потери, обусловленные резистивностью и неравномерным распределением тока, могут быть достаточно большими. Кроме того, за счет окружающих паразитных емкостей могут возникать нежелательные резонансы. По сравнению с другими дискретными компонентами добротность спиральных индуктивностей оказывается слишком малой [56].

5.3.2.3. Умножители напряжения

В ситуациях, когда появляется необходимо повысить напряжение, применяются умножители напряжения. Умножители напряжения необходимы тогда, когда используется память EEPROM: для записи информации в этом случае требуется повышенное напряжение. Умножители создаются каскадным соединением умножающих секций из диодов и конденсаторов. При этом повышение выходного напряжения осуществляется за счет суммирования напряжений заряженных секций.

5.3.2.4. Диоды

В метках RFID диоды необходимы как для преобразования высокочастотной энергии в энергию постоянного тока – как выпрямляющие элементы, так и для детектирования сигнала. В качестве диодов могут использоваться транзисторы, однако они требуют большего смещения и имеют худшие частотные свойства. Чаще всего, особенно на высоких частотах, предпочтительно использование.

Диоды с барьером Шоттки выполняются в виде перехода металл-полупроводник (р-типа или n-типа). В зависимости от конфигурации такие диоды могут иметь почти нулевое смещение, высокую частоту переключения, относительно низкую емкость перехода и малую потребляемую мощность. Они часто используются в высокочастотных детекторах, смесителях, переключательных схемах и выпрямителях. В комбинации с конденсаторами диоды Шоттки используются в удвоителях напряжения [54, 61].

5.3.2.5. Генераторы

Генераторы, по определению, предназначены для формирования сигналов чаще всего с постоянной частотой. В метках используют достаточно простые генераторы, которые, однако, обладают плохими характеристиками – нестабильной амплитудой и частотой. В зависимости от необходимых характеристик в метках RFID могут использоваться генераторы с достаточно хорошими техническими параметрами, что,

однако, достигается за счет их повышенной сложности, увеличенных размеров и, соответственно, большой стоимости.

5.3.2.6. Схемы ФАПЧ

Схемы, реализующие фазовую автоподстройку частоты (ФАПЧ), обеспечивают возможность синхронизации работы аппаратуры с фазой входного сигнала. Эти схемы часто используются в системах связи для выполнения различных функций. Они могут использоваться для программированной установки частоты в генераторах. Схемы ФАПЧ также применяются для частотной модуляции и демодуляции сигналов. В цифровых схемах они могут использоваться для восстановления синхронизации и генерации синхронизирующих сигналов [60].

В большинстве случаев схема ФАПЧ содержит фазовый детектор и генератор, управляемый напряжением, которые охвачены отрицательной обратной связью. Схемы ФАПЧ могут использоваться в метках RFID, однако из-за большого числа компонентов это приводит к избыточному усложнению, увеличению размеров и стоимости метки.

5.3.2.7. Цифровые схемы

Основным элементом построения вентиля является CMOS инвертор, который состоит из комбинации р-MOS и n-MOS транзисторов. Из этой пары транзисторов конструируется широкое многообразие логических схем, включая такие, как триггеры, регистры, мультивибраторы, сдвиговые регистры.

5.3.2.7.1. Триггеры и регистры

Триггеры являются логическими элементами, которые запоминают один бит двоичной информации. При помощи таких схем могут создаваться счетчики и другие более сложные устройства. Совместно с регистрами они часто используются для контроля вычислений в синхронных системах. В этом случае триггер может установить значение своего выходного сигнала равным значению входного сигнала синхронизирующей цепи. В простейшем случае триггер содержит два вентиля, каждый из которых обычно состоит из четырех транзисторов (две комплементарные пары р-MOS и n-MOS транзисторов).

Регистром обычно называют устройства для хранения определенного количества бит. Когда регистр находится в активном состоянии, выходной сигнал повторяет входной, а когда в пассивном – он фиксирует последнее значение сигнала. Регистр может быть активным во время одного такта и

пассивным – во время другого такта, когда он используется для контроля функционирования синхронных систем.

5.3.2.7.2. Сдвиговые регистры

Сдвиговый регистр образуется последовательным соединением триггеров таким образом, что выход одного соединен с входом другого, а все их установочные входы объединены. При подаче синхроимпульсов данные сдвигаются в одну сторону, в то время как данные поступают на его вход с другой стороны. Сдвиговые регистры удобны для преобразования параллельных данных в последовательные данные, и наоборот. Они также могут использоваться для временного хранения данных.

5.3.2.8. Мультивибраторы

Мультивибраторы являются разновидностью триггеров, у которых выход одного из вентиля соединен с входом другого вентиля через конденсатор. Эти устройства остаются в одном состоянии до тех пор, пока под воздействием импульса они не переходят в другое состояние. После некоторой задержки, время которой определяется параметрами схемы, мультивибраторы возвращаются в прежнее состояние. Они меняют свое состояние фронтом или спадом импульса и часто используются для генерации импульсов переменной длительности. Из-за того, что мультивибраторы строятся на основе аналоговых и цифровых элементов, они нестабильны и в большинстве случаев требуют дублирования [62].

5.3.2.9. Схемы памяти

Схемы памяти в интегральном виде могут реализовываться различными способами. Для применений в системах RFID низкой стоимости, когда метка просто хранит уникальный идентификационный номер (Unique Identification Number – UID), требования к схемам памяти невысоки. Практически единственным требованием является минимальная занимаемая площадь. Кратко рассмотрим некоторые из основных типов схем памяти, предназначенных для хранения UID: ROM, WORM, EEPROM и FRAM.

Данные в только читаемую память (Read Only Memory – ROM) обычно вводят в процессе производства при помощи лазера или напрямую - маской. Однажды записываемая и многократно читаемая память (Write Only Read Many – WORM) позволяет пользователю один раз записывать данные в схему памяти. Электрически перепрограммируемая память (Electrically Erasable Read Only Memory – EEPROM) позволяет многократно как записывать, так и считывать информацию, однако при этом она требует

высокого напряжения и имеет относительно большую потребляемую мощность. По сравнению с другими типами такая память обычно имеет большую площадь и время записи. Дальнейшим развитием EEPROM памяти является ферроэлектрическая память со случайным доступом – FRAM. Так же, как EEPROM, такая память может хранить данные в течение длительного времени, но процесс записи не требует высокого напряжения и имеет значительно меньшую мощность потребления [6].

5.3.3. Энергопотребление интегральных схем

Понимание важности характеристик энергопотребления интегральных схем в метке существенно не только с точки зрения проектирования самих схем, но и с точки зрения разработки протоколов связи. Кодирование, модуляция и команды могут и должны разрабатываться с учетом характеристик потребляемой мощности. Минимизация потребляемой мощности пассивной метки RFID непосредственно связана с повышением дальности считывания.

В метке используются как аналоговые, так и цифровые элементы. Несмотря на то, что цифровых цепей зачастую больше, чем аналоговых, энергопотребление аналоговых цепей непропорционально велико. По этой причине аналоговые цепи должны проектироваться наиболее грамотно с точки зрения потребляемой мощности [63]. Согласующие цепи антенн, входные цепи и источник питания должны разрабатываться с учетом максимального коэффициента передачи по мощности.

CMOS цифровые схемы имеют наилучшие характеристики энергопотребления, однако и вклад этих цепей в общее потребление мощности чипа очень существенен. Энергопотребление цифровых CMOS схем подразделяется на статическое и динамическое. Статической мощностью потребления называется мощность, которую потребляет схема в состоянии покоя. В идеале она равна нулю. Существуют паразитные утечки в диодах и других элементах, но в большинстве случаев ими можно пренебречь. В основном энергопотребление цифровых CMOS схем определяется динамическим потреблением. Кратко рассмотрим динамическое энергопотребление.

Динамическое энергопотребление состоит из двух составляющих: потребление короткого замыкания и емкостное потребление. Мощность потребления короткого замыкания обусловлена кратковременным протеканием тока от источника питания на землю в момент переключения вентилей. Можно показать, что в типичных случаях эта компонента составляет около 10% от емкостной составляющей, поэтому чаще всего ее не учитывают [64].

Емкостная составляющая мощности потребления обусловлена перезарядом паразитных емкостей интегральной схемы. Из предыдущего материала видно, что электронные схемы и цепи их соединений имеют множество паразитных емкостей. Энергия заряда емкости зависит от приложенного напряжения, поэтому общая мощность потребления связана с активностью и частотой переключения. Пренебрегая составляющей потребления короткого замыкания, рассмотрим формулу динамического потребления мощности:

$$P_{dyn} = \alpha f C V^2, \quad (5.2)$$

где α – коэффициент активности, f – скорость передачи данных, C – суммарная паразитная емкость, а V – напряжение питания. Коэффициент активности α представляет собой ожидаемое число переходов от 0 к 1 за время передачи данных. Средняя скорость передачи данных f является частотой синхронизации системы. Таким образом, динамическое потребление мощности в основном определяется тремя основными параметрами: активностью переключений, емкостью и квадратом напряжения питания.

Благодаря квадратичной зависимости согласно (5.2) уменьшение напряжения питания приводит к наибольшему снижению энергопотребления метки. В принципе снижение энергопотребления является очень важным системным фактором. Основная проблема при этом проявляется в том, что уменьшение напряжения питания приводит к возрастанию времени задержки метки, что снижает быстродействие системы RFID в целом. Возможно, однако, имеет смысл уменьшать напряжение за счет увеличения емкости и активности переключений [65].

Паразитные емкости присутствуют и в элементах схемы и в соединениях между ними. Эти емкости могут быть уменьшены за счет сокращения ширины соединительных линий, размеров элементов схем и расстояний между ними. При этом не следует забывать о краевых эффектах, которые приводят к уменьшению токов заряда емкостей, что может приводить к снижению быстродействия и, следовательно, к необходимости повышения напряжения питания.

Третьим параметром, влияющим на потребляемую мощность, является коэффициент активности. Скорость изменения информационных данных f в сущности является частотой переключений. В синхронных системах величина f соответствует частоте синхронизации. Без учета импульсных помех (glitching) коэффициент активности α является вероятностью, с которой происходят переходы уровней в течение одного бита данных. Высокочастотные импульсные помехи могут привести к тому, что α

превысит 1. Грамотное проектирование минимизирует такие ситуации. Можно показать, что для случайных сигналов максимальная величина коэффициента активности α равна 0,5; для простых логических вентилях она составляет от 0,4 до 0,5, а для конечных автоматов – находится в пределах от 0,08 до 0,18 [66].

Таким образом, можно достичь снижения потребляемой мощности за счет уменьшения величины перечисленных параметров. Другой способ снижения мощности состоит в параллельном выполнении определенных функций схемы. Это может привести к увеличению ее площади, но в итоге позволяет улучшить характеристики при сниженном напряжении. Другой практический способ состоит в отказе от бесполезных энергозатрат, таких как работа синхронизирующего генератора во время пассивного состояния аппаратуры или отказ от ее избыточных характеристик [65]. Возможно, также использование энергосберегающих CMOS элементов, способных сохранять энергию, которая обычно рассеивается в виде тепла [67].

5.4. Соединительные цепи и корпусирование

Необходимо также учитывать вопросы присоединения интегральной схемы (чипа) к антенне метки и корпусирования метки в целом. Соединительные цепи и корпусирование влияют не только на габариты и стоимость, но и на общие характеристики аппаратуры. Соединительные цепи влияют на эффективность передачи мощности между антенной и интегральной схемой. Корпусирование влияет не только на надежность и устойчивость, но и на чувствительность метки к окружающей среде и к свойствам маркируемых объектов.

Таким образом, варианты исполнения соединительных цепей и корпусирования метки выбираются в результате компромисса между стоимостью, которая, в свою очередь, связана с выбором материалов и технологией производства, и техническими характеристиками.

5.5. Резюме

В этом разделе мы вновь вернулись к некоторым основным вопросам, которые связаны с физическим аппаратным исполнением метки. Антенна, интегральная схема и корпусирование - все это может существенно отражаться на размерах, стоимости и остальных характеристиках метки. При проектировании метки особенно важен выбор материалов и геометрия антенн. Причем и то и другое существенно влияет на ее стоимость. С точки зрения интегральной схемы для того, чтобы снизить стоимость, размеры чипа должны быть минимальными. Для использования в метках RFID наилучшим образом подходит CMOS технология, которая в настоящее

время используется наиболее широко при проектировании и производстве микросхем. Доминирующее влияние на энергопотребление CMOS элементов оказывают такие параметры, как активность переключений, паразитные емкости и напряжение питания. Для достижения максимальной дальности считывания эти параметры необходимо минимизировать.

После краткого рассмотрения аппаратных вопросов, относящихся к метке, в следующей главе мы рассмотрим вопросы программного обеспечения.

Глава 6

Протоколы команд

6.1. Введение

После рассмотрения электродинамических, связных, административных и аппаратурных ограничений в технологии RFID, необходимо обратиться к алгоритмам и командным протоколам верхнего уровня, которые собственно осуществляют идентификацию меток.

Поскольку метка является простым носителем идентификационного номера, появляется задача точного чтения этого номера. Если в рабочей зоне считывателя находится единственная метка, не требуется никаких команд. При достаточной энергетике метка просто передает свои содержащиеся в ней данные. Однако, если в рабочей зоне считывателя находится множество меток, отвечающих одновременно, их сигналы интерферируют. Такое наложение сигналов в нашем случае называется коллизией, а результаты считывания чаще всего оказываются потерянными. Для избежания коллизий система RFID требует формирования команд, основанных на некоторых протоколах. Такие протоколы обычно называют антиколлизийными протоколами или алгоритмами.

Для того чтобы выбрать необходимый антиколлизийный протокол и команды, следует рассмотреть системные ограничения. Особое внимание следует обратить на надежность коммуникации (связи) и ширину спектра, а также на технические параметры чипа метки, которые непосредственно влияют на потребляемую мощность и стоимость. В первую очередь необходимо сужать ширину спектра, повышать надежность связи и минимизировать требования к чипу метки.

Для того чтобы успешно взаимодействовать с одной меткой среди множества меток, находящихся в поле считывания, необходимо реализовать эксклюзивный канал связи с этой меткой. Для этого существует достаточное количество решений – алгоритмов. После обзора этих алгоритмов мы проанализируем реализацию и ограничения двух наиболее распространенных из них. После этого мы обсудим основные командные требования, связанные с этими алгоритмами.

6.2. Обзор антиколлизийных алгоритмов

По мере увеличения числа маркируемых объектов вероятность одновременно считываемых меток увеличивается. Соответственно возрастает вероятность коллизий сигналов. Поэтому и в настоящее время

существует и в будущем возрастает необходимость в эффективных антиколлизийных алгоритмах.

Антиколлизийные алгоритмы, используемые в системах RFID, сходны со способами разрешения конфликтных ситуаций множественного коммуникационного доступа и с различными сетевыми протоколами, включая протоколы Aloha и семейство протоколов CSMA (Carrier Sense Multiple Access) [68]. Реализация антиколлизийных алгоритмов в технологии RFID, однако, ограничена низкой потребляемой мощностью и малым объемом, а иногда и полным отсутствием памяти метки. Кроме того, алгоритмы должны быть оптимизированы с учетом малого энергопотребления меток, чтобы не снижать дальность в случае пассивных, или увеличивать срок службы элемента питания в случае активных меток. К тому же из-за того, что метки способны взаимодействовать только со считывателем, использование методов CSMA невозможно. Более того, вариации параметров канала распространения сигнала в беспроводных каналах связи значительно больше аналогичных вариаций в проводных каналах – импульсные шумы крайне неблагоприятно влияют на относительно кратковременные сеансы связи между считывателем и меткой. Сложность алгоритмов также должна быть минимизирована, так как это приводит к увеличению стоимости аппаратуры.

Существует большое разнообразие антиколлизийных методов, которые могут классифицироваться различными способами [69]. Наиболее общими классификационными признаками являются: пространство, частота и время (рис.6.1).

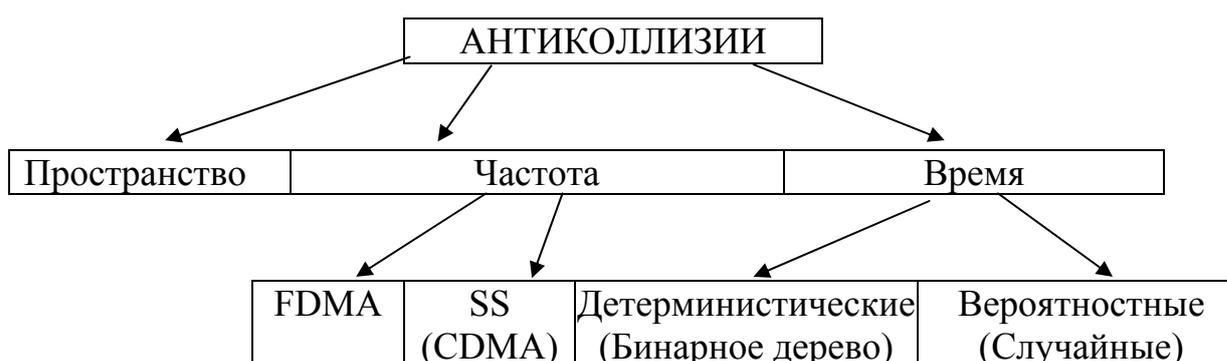


Рис.6.1. Классификация антиколлизийных алгоритмов.

В случае пространственных алгоритмов метки обычно локализуются в пространстве таким образом, чтобы обеспечить их последовательное считывание. Это достигается изменением зоны считывания или, в случае пассивных меток, вариацией мощности, излучаемой считывателем. Известен способ, использующий обе эти возможности на основе определения

максимального отклика меток на различных расстояниях [70]. Другой способ использует ряд считывателей с малой дальностью действия, такой, которая обеспечивает считывание только одной метки в поле каждого считывателя в данное время [6]. Еще в одном способе разделения меток и их последовательного считывания предлагается триангуляционное использование сверхширокополосной связи с определением местоположения [71]. Основной проблемой пространственных методов является сложность достижения высокой точности определения дальности. Требования к точности определения дальности еще более возрастают с увеличением числа меток в зоне действия считывателя и, соответственно, уменьшением расстояния между ними. В настоящее время считается, что наилучшим является использование пространственных методов в сочетании с частотными и временными методами.

Для обеспечения уверенной беспроводной связи обычно используются частотные методы. Системы FDMA (Frequency Domain Multiple Access) используют разделение общей полосы частот на фиксированное число каналов. В системах RFID низкой стоимости такое решение не применимо, так как при этом потребуются высокостабильные генераторы и селективные полосовые фильтры. Технология Magellan использует комбинацию FDMA и TDMA (Time Domain Multiple Access) методов. Поскольку достоинства такой системы недостаточно обоснованы, не очевидно как это повлияет на технические характеристики и стоимость системы RFID.

Системы CDMA (Code Domain Multiple Access) имеют много преимуществ перед FDMA системами, так как они лучше адаптированы к изменению трафика, имеют большую емкость и простое управление процессом [68]. Системы, основанные на применении CDMA и SS (Spread Spectrum) методов достаточно сложны и дороги. К тому же их использование может ограничиваться частотными ресурсами, предусмотренными в регламентах. Поэтому SS методы, включая FH (Frequency Hopping) и DS (Direct Sequence), могут быть реализованы только в UHF или микроволновом диапазоне, где имеются соответствующие частотные ресурсы.

подавляющее число антиколлизийных алгоритмов в технологии RFID основано на использовании временных методов, в которых момент передачи сигнала изменяется во времени. Эти алгоритмы подразделяются на детерминистические и вероятностные.

Детерминистический алгоритм реализуется, когда считыватель генерирует запрос или команду, которая возбуждает определенную метку с уникальным идентификационным номером UID (Unique Identification Number). На основании этого номера считыватель или перебирает список известных номеров, или выполняет определенные действия по поиску бинарным способом. Переборные методы особенно эффективны, когда в

зоне действия считывателя находится небольшое количество меток. При этом также требуется предварительное знание всех номеров меток.

В настоящее время наиболее широкое применение находят бинарные алгоритмы. Компания TI (Texas Instrument) использует бинарный алгоритм [72]. Существуют различные варианты такого алгоритма. Компания Jacomet и др. [74] предложила относительный алгоритм бинарного поиска. Компания SCS предложила другой вариант этого алгоритма [75]. Некоторые алгоритмы бинарного поиска могут быть достаточно быстродействующими, однако все же они работают достаточно корректно, если в течение поиска в поле считывания не появляются дополнительные метки. Мы рассмотрим алгоритм, основанный на основе протокола Law.

Вероятностные алгоритмы - это такие методы разрешения коллизий, когда метки в поле считывателя генерируют сигналы в случайные моменты времени. Существует большое число решений, когда считыватель различным образом управляет метками. Значительное число способов основывается на протоколе Aloha, предназначенном для множественного сетевого доступа [68]. По этой схеме узел передает пакет после приема пакета. В случае возникновения коллизии узел входит в насыщение и передает пакет снова после случайной задержки. В стробированном варианте протокола Aloha (slotted Aloha) передача пакета производится в течение фиксированного периода, большего некоторого известного постоянного времени. В патенте Furuta [76] описывается вариант такого протокола для бесконтактных персональных карт. Международный стандарт ISO 15693 поддерживает метод, аналогичный стробированному варианту протокола Aloha.

Алгоритм SuperTag компании BTG [77] функционирует по принципу дополненного стробированного варианта алгоритма Aloha. После приема данных метки могут прекратить работу или передавать данные с пониженной частотой повторения. В то время, когда метки не работают, алгоритм SuperTag пересчитывает идентично кодированные объекты. Для того чтобы быть уверенным, что при передаче данных не случится коллизия, другой вариант протокола SuperTag предписывает выключение всех меток, кроме одной. После определенного периода времени метки активизируются и цикл повторяется. По другим предложенным способам считыватель работает с перерывами – в импульсном режиме, предписывая работу меток после случайной задержки [78, 79]. После того как метка передаст данные, она перестает работать, чтобы уменьшить вероятность будущих коллизий. Некоторые алгоритмы предписывают, чтобы по стартовой команде считывателя метки после случайной задержки отвечали не полным, а укороченным сигналом. После этого считыватель предписывает меткам с большей задержкой замолчать, а одной метке передавать полный пакет данных [80]. В качестве типичного представителя таких протоколов мы анализируем семейство алгоритмов SuperTag.

Многие антиколлизийные алгоритмы требуют обнаружения факта возникновения коллизии сигналов. Наиболее общепринятый метод обнаружения коллизий основан на использовании свойств кодирующих сигналов. Код NRZ и другие коды [6], связанные лишь с уровнем сигнала, принципиально не пригодны для обнаружения коллизий (рис. 6.2a). В то же время код Манчестера и другие коды, в которых информация связана с переходом сигнала от одного уровня к другому (рис. 6.2b), обладают такой возможностью.

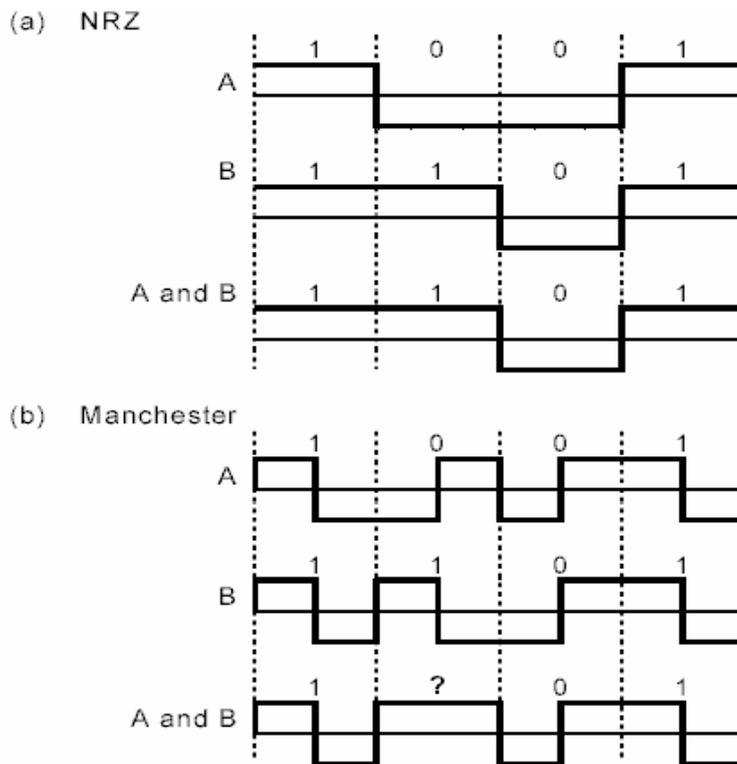


Рис. 6.2. Обнаружение коллизий с использованием различных способов кодирования сигналов.

Известны и другие методы обнаружения коллизий, которые основаны на использовании модуляционных схем. Считыватель по «дрожанию» (wobbling) сигнала способен обнаружить одновременный прием нескольких меток, которые используют FSK модуляцию своих сигналов [78].

6.2.1. Анализ антиколлизийных алгоритмов

В этом разделе мы проанализируем два протокола с точки зрения их исполнения и характеристик. Несмотря на то, что все антиколлизийные алгоритмы уникальны и имеют свои достоинства и недостатки, достаточно рассмотреть два семейства алгоритмов – SuperTag и QT, которые, в определенном смысле, являются их типичными представителями. Оба алгоритма реализуются во временной области. Однако алгоритм SuperTag является вероятностным, а QT – детерминистическим.

6.2.1.1. Алгоритм SuperTag

Алгоритм SuperTag был разработан южноафриканской компанией SCIR и в настоящее время лицензирован компанией BTG. Он представляет собой семейство алгоритмов, основанных на сетевом протоколе Aloha. Для того чтобы обеспечить идентификацию множества меток, как упоминалось ранее, семейство алгоритмов SuperTag использует вероятностный подход. Известны четыре варианта алгоритма (рис. 6.3). Каждый вариант подразумевает использование стартовой команды считывателя [77].



Рис. 6.3. Четыре варианта алгоритма SuperTag. Черный цвет соответствует успешному, а серый – неуспешному разрешению коллизий.

Tag - Метка

В простейшем варианте после приема стартовой команды метки со случайной задержкой отвечают полным идентификационным номером. В соответствии с этой случайной задержкой метки продолжают отвечать даже после идентификации. Это соответствует варианту ST.std.free (standard free-running) алгоритма SuperTag.

В несколько более сложном варианте алгоритма ST.std.off (standard shut-off) считыватель после приема полного номера метки генерирует команду на прекращение ее ответа.

В варианте алгоритма ST.std.free метки могут начать ответ во время ответа других меток, а это вызывает коллизии. От этого недостатка свободен другой вариант, который при помощи генерируемой считывателем команды запрещает ответы всех меток, кроме первой ответившей. Этот вариант соответствует алгоритму ST.fast.free.

Последний вариант алгоритма SuperTag наиболее сложен. Как и в предыдущем варианте – ST.fast.free, считыватель запрещает ответы всех меток, кроме первой ответившей. Однако, в дополнение, считыватель выключает метку сразу после ее идентификации. Этот вариант соответствует алгоритму ST.fast.off.

6.2.1.2. Бинарные алгоритмы с запросом

Другой широкий класс антиколлизийных алгоритмов во временной области представляет собой детерминистические протоколы, которые для определения уникального номера UID используют бинарный поиск. Эти протоколы классифицируют в зависимости от того, какая информация требуется при передаче сигнала метки. Некоторые алгоритмы требуют, чтобы метки в течение поиска отвечали полным или почти полным номером UID. Другие алгоритмы заставляют считыватель достраивать UID бит за битом, при этом метки просто отвечают при совпадении запроса. Мы остановимся на этом позже, так как ситуация может упрощаться исходя из требований к метке. Многие особенности предложенных методов [73, 75] мы рассмотрим на основе анализа алгоритмов QT (рис.6.4).

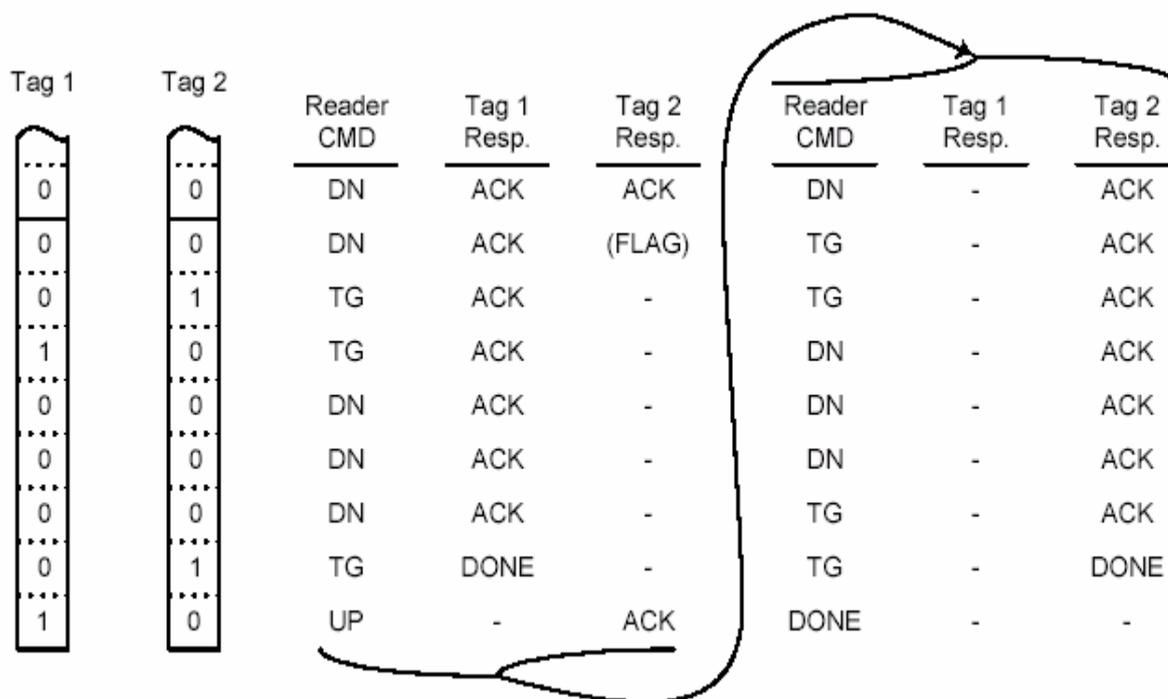


Рис.6.4. Последние девять бит выборки QT.ds цикла в присутствии двух отвечающих меток.

Tag - Метка, Reader - Считыватель, Resp.- Ответ.

По этой схеме считыватель может генерировать две команды поиска: одна – DN, инструктирует метку двигаться по дереву, другая – TG (toggle), инструктирует метку перейти на другую ветвь дерева и двигаться по ней. Если последний идентифицируемый бит был равен 0, команда DN инструктирует метку двигаться дальше по ветви 0. Если последний бит равен 1, считыватель инструктирует метку далее двигаться по ветви 1. Команда перехода TG инструктирует метку перейти на ветвь, которая противоположна значению последнего бита. Если в следующей битовой позиции метка имеют совпадение с 0 или 1, она отвечает подтверждением. В случае несовпадения битовых значений метка перестает отвечать и битовый указатель останавливается на последнем бите, который был удачно идентифицирован.

После генерации стартовой команды в битовом указателе метки устанавливается 0 и считыватель может генерировать DN или TG команду. Считыватель в зависимости от конкретного алгоритма может генерировать последовательность DN или TG команд. После генерации каждой команды считыватель анализирует подтверждения меток о битовом совпадении. Это продолжается до тех пор, пока в результате анализа будет получен полный 64-битный UID и одна метка будет идентифицирована. После совпадения номера метка генерирует подтверждение и выключается. В предыдущих

рассуждениях мы предполагали, что команды DN генерируются после TG команд, если не поступает подтверждения совпадения.

Как отмечалось ранее, если в метке нет совпадения с 0 или 1 в текущей битовой позиции, она прекращает отвечать и битовый указатель останавливается на последнем идентифицированном бите. Кроме того, метка устанавливает флаг, обозначая ошибку. После того, как путь по дереву завершен и одна метка идентифицирована, считыватель может инициировать UP команду. Эта команда предписывает меткам снять флажки и установить их указатели на ближайшие 8 бит, привязанные к полным 64 битам UID. Если на данной ветви дерева метки отсутствуют, может быть инициирована еще одна команда UP. Такой алгоритм называется QT.ds.

В рассмотренных алгоритмах, основанных на бинарном поиске, распределение номеров UID влияет на характеристики аппаратуры. Наилучшие характеристики обеспечиваются при последовательном распределении номеров. Наихудшие характеристики обеспечиваются при случайном распределении номеров. Мы рассмотрим оба случая.

6.2.1.3. Моделирование

Результаты моделирования показаны в табл. 6.1.

Число меток в поле	Время разрешения коллизий, (с)					
	QT.ds.(посл.)	QT.ds.(случ.)	ST.std.free	ST.std.off	ST.fast.free	ST.fast.off
10	0.0218	0.117	0.18	0.09	0.038	0.021
50	0.07	0.656	2.8	0.807	0.447	0.183
100	0.13	1.21	5.22	2.316	1.035	0.326
500	0.922	7.32	>200	17.314	6.844	1.683
1000	4.01	15		>200	18.734	3.945

Таблица 6.1. Результаты моделирования времени, необходимого для считывания различного количества меток в поле считывания.

Данные по алгоритму Supertag получены моделированием CSIR SuperSim. Результаты получены в результате усреднения 10 попыток для различного числа меток, при этом в качестве оптимального выбрано максимальное время. Также в результате моделирования получены данные по алгоритму QT.ds. В каждом случае скорость обмена принималась равной 64 кб/с, а длина идентификационного кода – 64 бит. Несмотря на то, что данные получены посредством небольшой выборки и доверительный интервал велик, приведенные в табл. 6.1 результаты моделирования позволяют оценить основные тенденции при использовании типичных антиколлизийных алгоритмов.

6.2.1.4. Сравнение результатов моделирования

Два типа алгоритмов, рассмотренных нами, могут сравниваться по двум критериям: технические характеристики и практическая реализация. Что касается технических характеристик, необходимо обратить особенное внимание на два фактора – скорость идентификации и поведение аппаратуры при изменении числа меток в процессе считывания. Также следует учитывать характеристики в присутствии шума и другие специфические достоинства и недостатки. Что касается практической реализации, то особенно важной является аппаратная реализация и вопросы программного обеспечения (команды). На основе данных, приведенных в табл. 6.1, по этим параметрам мы и сравним алгоритмы SuperTag и QT.

При рассмотрении алгоритмов SuperTag можно заметить, что при переходе от варианта ST.std.free к варианту ST.std.off, наблюдается близкое к линейному возрастание времени разрешения коллизий в зависимости от числа меток. Это обусловлено дополнительными функциями варианта ST.std.free, которые выключают метки после их идентификации и/или выключают метки, когда одна из них отвечает. Обе эти функции приводят к уменьшению числа отвечающих в поле меток и, следовательно, имеют линейную тенденцию. Видно также, что быстродействие алгоритма ST.fast.off в зависимости от числа меток в поле близко к линейному.

Обращаясь к алгоритму QT.ds мы видим существенные различия в характеристиках при последовательном и случайном распределении номеров меток. В определенных применениях – на производстве или в дистрибьюторских центрах, например, номера UID могут быть распределены последовательно. В других применениях, таких как розничные магазины, номера распределяются случайно. Мы рассмотрим оба случая. В алгоритмах SuperTag связь с распределением номеров меток отсутствует. Мы также видим, что быстродействие алгоритма QT.ds по отношению к числу меток в поле при случайном распределении номеров близко к линейному, а при последовательном распределении – имеется резко выраженная нелинейность. Кроме того, как только число меток в поле приближается к «групповой» емкости в 256 меток (благодаря 8-битной команде UP), скорость считывания меток возрастает. Это, вероятно, объясняет нелинейности, видимые при последовательном распределении номеров UID.

Сравнивая характеристики двух алгоритмов, заметим, что они близки при малом числе меток в поле. При увеличении числа меток в поле проявляются нелинейности.

Теперь оценим количество команд, которое требуется передавать считывателю для разрешения коллизий. Из всех рассмотренных алгоритмов, вероятно, алгоритм ST.fast.free требует наименьшего числа команд от

считывателя. Это может быть очень полезным в тех частотных диапазонах, где действуют жесткие регламентные ограничения. Прямой поиск алгоритма QT.ds, как и в случае ST.fast.off, требует передачи значительного числа команд. В общем случае, чем больше объем передаваемых команд, тем шире требуемая полоса частот и тем выше вероятность возникновения ошибок.

Кроме технических характеристик и команд, важна практическая аппаратная реализация. Алгоритмы SuperTag требуют генерации случайных задержек перед ответами. А это требует соответствующих схемных решений. Сложность реализации вариантов различна. Некоторые решения для генерации случайных чисел предполагают использование собственных свойств электронных схем [80]. В других решениях используют специальные генераторные схемы случайных чисел. Усовершенствованные алгоритмы SuperTag должны также обладать способностью распознавания команд выключения и прекращения ответа. Алгоритмы QT.ds. требуют наличия нескольких компараторов для сравнения передаваемых бит с битами, хранящимися в памяти. Они также должны устанавливать флаги битам и некоторым состояниям информации. В любом случае избыточность в схемных решениях требует дополнительной площади и, соответственно, приводит к большей стоимости чипа.

Каждый из алгоритмов – SuperTag и QT.ds, имеет свои уникальные достоинства и недостатки. Так алгоритм QT.ds имеет присущую ему способность селективировать метки с определенными номерами. Это может быть удобным в применениях, когда должны распознаваться или наоборот, не распознаваться, определенные, специфические метки или группы меток.

В свою очередь алгоритмы SuperTag, по крайней мере, ST.std.off, требуют меньшего числа команд считывателя, более узкой полосы частот и, следовательно, можно предположить наименьшую вероятность ошибок.

Уникальным недостатком алгоритма QT.ds и других протоколов с бинарным поиском является их неспособность во время поиска идентифицировать метки, которые вновь поступают в поле считывания. Эта особенность может быть важной или не важной в зависимости от конкретного применения аппаратуры RFID.

Проблема алгоритмов SuperTag и других, связанных со случайным временем ответа, состоит в необходимости назначать максимальное время задержки ответа метки. Это время может передаваться метке или быть заложенным в производстве.

6.2.2. Локализация

В зависимости от аппаратной реализации и рабочего диапазона частот антиколлизийные схемы могут быть локализованы в частотной, временной или пространственной областях или в комбинации этих областей. Такая комбинация может повысить скорость идентификации. В то же время это

может усложнить систему RFID или увеличить вероятность появления ошибок, то есть понизить достоверность идентификации.

6.2.3. Системные вопросы

Характеристики индивидуальных идентификационных алгоритмов несомненно важны, однако наиболее важным этапом проектирования является общесистемный уровень. Грамотный выбор способа кодирования и алгоритма разрешения коллизий, который обеспечит распознавание отдельных кодов или групп кодов, может позволить перенести значительную часть сложного антиколлизийного процесса из метки в считыватель или хост. Например, в определенных ситуациях большая часть идентификационного кода может быть известна или может потребоваться идентификация меток только с определенными кодами. В таких конкретных ситуациях считыватель может напрямую адресовать информацию о метках с такими кодами в хост. Детерминистические алгоритмы позволяют реализовать такую возможность, а вероятностные – нет.

6.3. Команды

С точки зрения протоколов и команд существует два основных класса систем RFID: метка говорит (передает данные) первой (ТТФ – Tag Talk First)) и считыватель говорит (передает команды) первым (РТФ – Reader Talk First)). В ТТФ системах метки отвечают сразу, как только обеспечиваются энергией. В РТФ системе даже при достаточной энергии метки находятся в режиме молчания. Они отвечают только по команде считывателя. Методы РТФ предпочтительны в большинстве случаев, так как при этом исключается интерференция сигналов от несовместимых меток. Например, если в поле считывания находятся метки различных поставщиков, благодаря различным стартовым командам ответят метки, совместимые только со своим считывателем. В ТТФ системе, все метки ответят одновременно, а их сигналы будут интерферировать и, возможно, создавать друг другу помехи.

Команды, как и антиколлизийные алгоритмы, характеризуются техническими параметрами и исполнением. Большой объем команд может обеспечивать лучшие характеристики, однако, может потребовать большего объема аппаратного обеспечения. Набор команд должен включать, по крайней мере, стартовую команду и команды, необходимые для запуска антиколлизийного алгоритма и достижения необходимых характеристик. Особенно важными характеристиками являются скорость идентификации и надежность связи.

Для расширения функциональности и сферы использования могут потребоваться дополнительные команды. В любом случае для только

читаемых меток предельно низкой стоимости требуется передача лишь идентификационного кода.

6.3.1. Команды и ширина спектра

При рассмотрении процессов кодирования и модуляции сигнала мы видели, что вероятность передачи компонента сигнала, скажем, нулевой амплитуды за тактовый период времени по сравнению с сигналом полной амплитуды в течение одной половины такта и нулевой амплитуды в течение другой половины такта, влияет на ширину спектра сигнала. Если вероятность передачи половины периода сигнала без модуляции выше, чем вероятность передачи половины периода с модуляцией, ширина спектра сигнала будет меньше. Этот эффект наиболее полезен для считывателя с точки зрения обеспечения необходимых санитарных и административных норм. Таким образом кодирование, модуляция и передача команд при максимально возможном уровне излучения, ограничиваются предельными нормами напряженности поля и ширины спектра. Это особенно должно учитываться в тех случаях, когда ширина спектра и связанная с ней мощность излучения резко лимитирована регламентами.

6.4. Резюме

В этой главе мы рассмотрели проблемы, относящиеся к протоколам команд и антиколлизийным алгоритмам. Так как целью пассивных систем RFID является простое считывание идентификационных кодов, которые хранятся в метке, а в поле считывания присутствует множество меток, протоколы команд должны оптимизироваться по отношению к антиколлизийным алгоритмам. При рассмотрении антиколлизийных алгоритмов важной характеристикой является не только скорость идентификации, но также требуемый набор команд, функционирование в присутствии шума и аппаратная реализация. Кроме того, с точки зрения команд и антиколлизийных алгоритмов очень важным оказывается влияние ширины спектра.

Глава 7

Влияние ограничений на технические параметры

7.1. Введение

Мы анализировали различные ограничения на системы RFID и определили основные факторы, которые влияют на аппаратуру и ее характеристики. В этой главе мы попытаемся рассмотреть совокупность этих ограничений таким образом, чтобы показать, как они взаимодействуют и влияют друг на друга, а также как они определяют технические характеристики, габариты и стоимость.

Основными техническими характеристиками систем RFID, которые мы выделили в главе 1, являются дальность, быстродействие, надежность связи и электромагнитная совместимость. Каждая отдельная характеристика аппаратуры непосредственно или косвенно определяется совокупностью всех ограничений.

Стоимость и габариты являются наиважнейшими характеристиками, которые, в сущности, очевидны, когда рассматривается совокупность технических характеристик при наличии всех ограничений. Многие проблемы, относящиеся к стоимости и габаритам аппаратуры, были рассмотрены в главе 6, однако мы еще обратимся к этому вопросу в дальнейшем.

7.2. Дальность

Дальность действия системы RFID в основном ограничивается напряженностью излучаемого электромагнитного поля и электродинамическими свойствами распространения сигнала в окружающей среде. Напряженность излучаемого поля непосредственно ограничена административными нормами (регламентами), а также косвенно, через административные нормы и ограничения по ширине спектра сигнала, определяется закономерностями процесса связи. В дальнейшем мы, используя напряженность поля, пространственную ориентацию и расположение метки относительно считывателя, а также влияние окружающего пространства, определим, как метка принимает, распределяет и потребляет мощность. В этом разделе мы рассмотрим ограничения и их влияние на дальность.

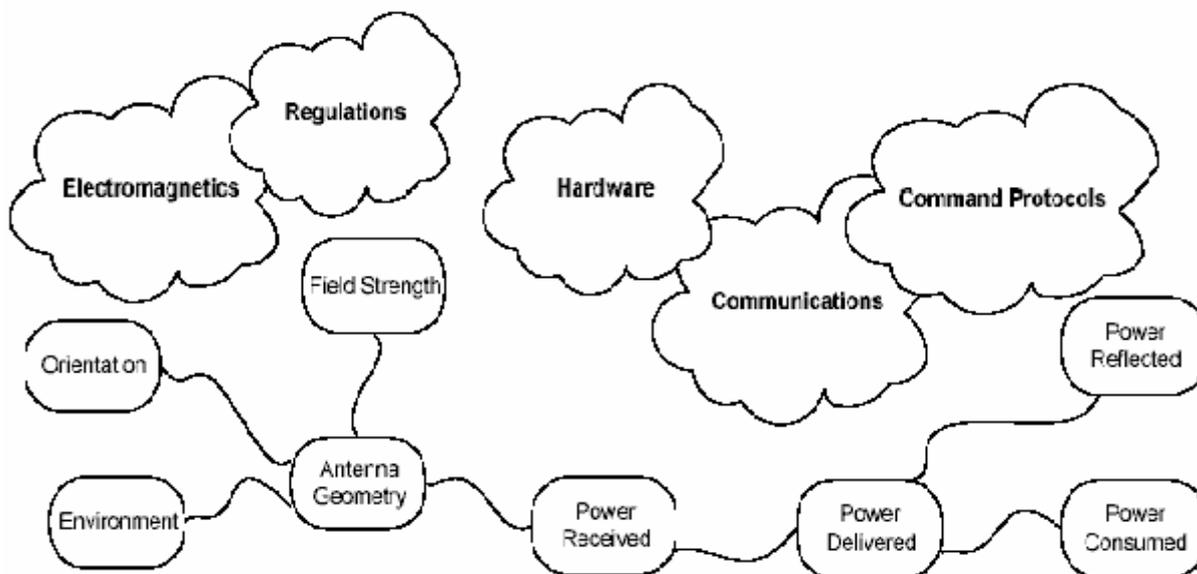


Рис.7.1. Основные факторы, влияющие на дальность.

Electromagnetics - Электродинамика, Regulations - Регламенты, Orientation - Ориентация, Field Strengths - Напряженность поля, Environment - Окружающее пространство, Antenna Geometry - Геометрия антенны, Power Received - Принятая мощность, Power Delivered - Доставленная мощность, Power Reflected - Отраженная мощность, Power Consumed - Потребляемая мощность, Hardware - Аппаратура, Communications - Связь, Command Protocols - Протоколы команд.

7.2.1. Напряженность поля, ориентация и окружающая среда

Дальность действия аппаратуры RFID часто определяется как максимально возможное значение дальности, однако такая ситуация имеет место только в ограниченном числе случаев. Чаще всего, на практике, местоположение и ориентация метки относительно считывателя произвольна. В главе 2 мы видели, что напряженность поля находится в большой зависимости от взаимной ориентации и расположения метки и считывателя. Напряженность поля очень существенно может изменяться в зависимости от вида антенны. Так, например, как в ближней, так и в дальней зонах поля могут использоваться высоконаправленные антенны. В ближней зоне для повышения направленности применяются ферриты. В дальней зоне высокая направленность антенн обеспечивается за счет использования решеток или других конструктивных решений. Хотя при заданной мощности применение направленных антенн и позволяет повысить дальность, все же достижимая дальность считывания ограничена. При этом в некоторых случаях использование высоконаправленных антенн необходимо, в других – требуются всенаправленные антенны.

Если регламенты допускают лишь предельную излучаемую мощность, то с целью повышения дальности могут использоваться высоконаправленные антенны. В типичных случаях, однако,

устанавливается предельная максимальная напряженность поля на определенном расстоянии или мощность на входе антенны и ее усиление. Так обстоит дело в большинстве администраций. Однако в некоторых администрациях на определенных частотах высоконаправленные антенны могут использоваться. Например, в США на частотах 902 МГц, 2,45 ГГц и 5,8 ГГц при использовании перестройки частоты предел излучаемой мощности установлен через входную мощность антенны и ее усиление. Также предписано, что усиление может превышать установленный уровень, но при этом входная мощность антенны должна быть пропорционально уменьшена, то есть устанавливается предел напряженности поля или плотности потока мощности на определенном расстоянии. Ситуация в Европе на упомянутых частотах похожа – тут пределы установлены в единицах EIRP или ERP. В Японии, однако, пределы обычно устанавливаются непосредственно для входной мощности и усиления антенны. Хотя эти параметры и связаны с EIRP, установленные пределы входной мощности и усиления антенны не должны превышать. В тех случаях, когда установлен предел EIRP, повышение направленности не позволяет увеличить дальность действия. В этом случае устанавливается размер зоны считывания, снижая при этом вероятность интерференции и уменьшая число меток в зоне и, тем самым, вероятно, облегчая разрешение коллизий сигналов. Формирование диаграмм направленности с контролируемым изменением напряженности поля и мощности (в установленных пределах) дает дополнительную степень свободы в управлении пространственным распределением поля с возможным облегчением разрешения коллизий.

Таким образом, мы видим, что административное регулирование ограничивает напряженность поля и плотность мощности в определенных областях пространства, даже если эти параметры могут обеспечиваться техническими средствами в широких пределах. Проблема состоит, однако, в том, что административные нормы устанавливают на определенном расстоянии максимальный уровень напряженности поля и плотности мощности. В то же время считыватель размещается в любой точке пространства в окружении проводящих предметов и его диаграмма направленности при этом может изменяться в очень широких пределах. Это влияние должно учитываться, когда рассматриваются специфические применения и конфигурации систем RFID.

7.2.2. Прием, распределение и потребление мощности сигнала

После того как определена напряженность поля или плотность мощности в некоторой точке пространства, следующей проблемой является прием мощности. Для того чтобы достичь максимального индуцируемого

напряжения или доступной мощности, необходимо оптимально ориентировать антенну метки по отношению к излучаемому полю. На практике это может быть очень сложной задачей. Вероятно максимально достижимая мощность может быть получена, если маркированный объект передвигается на конвейере и изменение его ориентации ограничено. Для многих применений, однако, такие условия не выполнимы. В зависимости от геометрии и характеристик антенна метки может быть более или менее чувствительна к ориентации. Кроме того, антенна считывателя, как и антенна метки, обладает собственной направленностью. Для уменьшения провалов приема могут использоваться специальные антенны, такие, как ортогональные рамки; однако в любом случае при выборе конфигурации системы и вариантов ее применения вопросы ориентации антенн должны рассматриваться особенно внимательно.

Напряженность поля и плотность потока мощности в некоторой точке пространства определяются регламентами и условиями распространения сигнала. Напряженность поля или плотность мощности воспринимается антенной метки с учетом ее геометрии и ориентации. При фиксированной геометрии и ориентации антенны функционирование метки зависит от параметров ее электронной схемы. Особенно важны импедансы антенны, согласующей цепи и нагрузки. В зависимости от индуцируемого напряжения они определяют ток, протекающий по цепям электронной схемы метки. В зависимости от рабочей частоты в ближней и дальней зонах поля для того, чтобы максимизировать индуцируемое напряжение, используются различные средства.

В ближней зоне и реактивной линии связи метка выступает в качестве стока энергии. При помощи резонанса в антеннах метки и считывателя метка, в некотором смысле, увеличивает пропускную способность стока. Обеспечение резонанса включением емкости параллельно катушке индуктивности метки или включение емкости последовательно катушке индуктивности считывателя позволяет достичь высокой добротности. Эту добротность определяет импеданс резонансной цепи и эквивалентный импеданс остальной части электронной схемы. Добротность определяется отношением запасенной к рассеиваемой энергии или отношением центральной частоты к полосе пропускания на уровне 3 дБ. Следовательно, благодаря добротности мы фактически имеем дело со схемой частотной фильтрации. Полосу пропускания этого фильтра необходимо соотносить с полосой пропускания антенны метки. При высокой добротности полоса пропускания может быть зауженной и должны быть приняты меры для того, чтобы не потерять информацию. Мы видели, что модуляция или перенос спектра на несущую частоту приводит к появлению боковых полос. Если их отфильтровать, то информация будет потеряна. Кроме того излишне высокая добротность приводит к сильному влиянию окружающей среды.

Окружающие антенну проводящие предметы, в том числе другие метки, изменяют резонансную частоту и, соответственно, добротность системы. В ситуации, когда вблизи одной метки находятся другие метки, выходом из положения может быть настройка этой метки на частоту выше номинальной.

В дальней зоне ситуация иная. Мощность, доставляемая в нагрузку, определяется степенью согласования импеданса антенны метки с эквивалентным импедансом электронной схемы. Если эти импедансы комплексно сопряжены, а это означает, что их реальные части равны, а реактивные части равны и противоположны по знаку, то обеспечивается максимальная мощность, доставляемая электронной схеме. При наличии расстройки часть мощности будет отражаться от электронной схемы и переизлучаться антенной в пространство. Собственно именно это явление и лежит в основе модуляции обратного рассеяния. Управление импедансом приводит к модуляции электромагнитной волны, падающей на метку. Чем большая часть сигнала отражается, тем меньшая его часть поглощается, и наоборот. При рассмотрении части мощности, поступающей в нагрузку, важна конфигурация самой нагрузки. Мощность, доставленная метке, расходуется не только на ее энергопитание, но и для обеспечения информационного обмена и синхронизации. Что касается конфигурации нагрузки, то выпрямительные диоды необходимы для обеспечения энергопитания и выделения информации. Энергия запасается конденсатором и доставляется цифровым схемам, которые реализуют процесс принятия решения. Решение принимается на основе анализа информации. Метка сравнивает принятую информацию с информацией, хранящейся в памяти, и принимает решение: передать хранящуюся информацию, не отвечать или выполнять иное действие. Все эти действия, естественно, требуют затрат мощности. Потребляемая мощность зависит от активности входящих в метку вентилей, паразитных емкостей электронной схемы и напряжения питания. Активность вентилей определяется частотой синхронизации и необходимостью выполнения антиколлизий операций. Вентильные схемы имеют определенный пороговый уровень напряжения, ниже которого они не функционируют. Мощность, доставленная к электронной схеме, должна обеспечивать этот необходимый уровень напряжения и быть достаточной для того, чтобы метка функционировала. Если метка не способна выполнять свои функции, ее следует переместить ближе к источнику излучения – считывателю, чтобы напряженность поля и доставленная мощность возросли.

Важно не только, чтобы мощность сигнала была достаточной для обеспечения энергопотребления метки; она также должна быть достаточной для обнаружения сигнала меткой. Если не будет достаточной разности уровней сигнала, уверенно воспринимаемых пороговой схемой метки, или не будет достаточного отношения сигнал/шум, обнаружение будет

неудовлетворительным. Следовательно необходимо такое расстояние между меткой и антенной считывателя, на котором обеспечивается достаточное соотношение уровней сигнала и шума, и которое позволяет достичь уверенной работы пороговой схемы метки и, следовательно, обеспечить заданную вероятность битовых ошибок BER.

До сих пор мы рассматривали прямую линию связи от считывателя к метке. Теперь рассмотрим обратную линию связи от метки к считывателю. В этой линии производится модуляция сигнала информацией, хранящейся в памяти метки. Это обеспечивается за счет вариации импеданса нагрузки или модуляции обратного рассеяния. Уровень модулированного таким образом сигнала значительно ниже уровня сигнала считывателя. Однако, благодаря сложной обработке сигнала в считывателе, это обычно не ограничивает дальность действия системы.

Таким образом, мы увидели, что на дальность влияет не один, а множество факторов. Эти факторы можно использовать. Негативные факторы могут быть минимизированы, а позитивные – подчеркнуты. При имеющихся ограничениях электродинамики, регламентов, связи, проектирования и реализации аппаратуры и антиколлизийных алгоритмов все же остается достаточное число степеней свободы. Дальность является параметром, который зависит от множества факторов. После обзора вопросов, касающихся дальности, необходимо перейти к рассмотрению практического использования, конфигурации системы и других вопросов технологии RFID.

7.3. Быстродействие и надежность связи

Следующим основным параметром системы RFID является

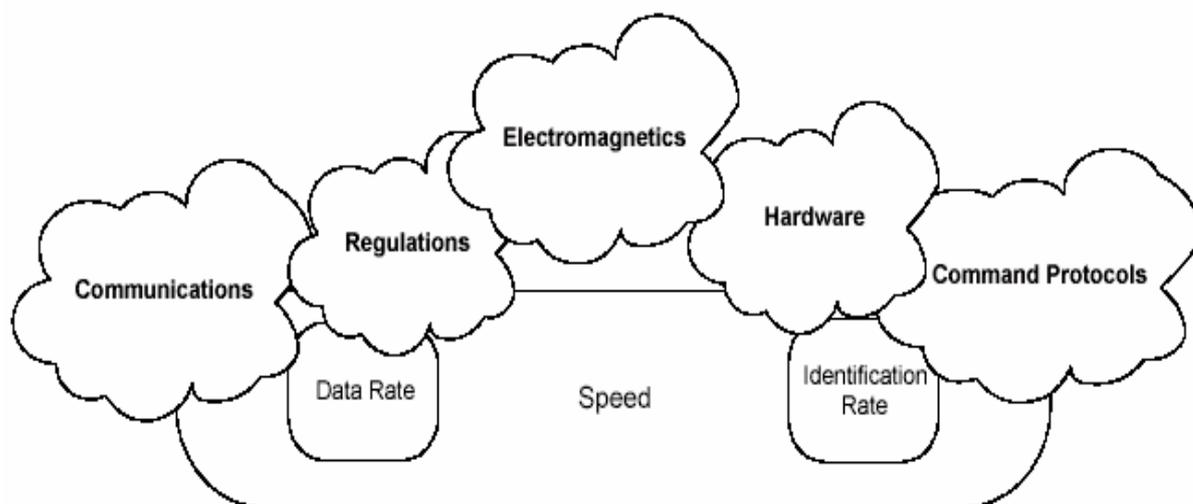


Рис. 7.2. Факторы, влияющие на быстродействие и ограничения, влияющие на них.

Communications - Связь, Regulations - Регламенты, Data Rate - Скорость передачи данных, Electromagnetics - Электродинамика, Speed - Быстродействие, Hardware - Аппаратура, Identification Rate - Быстродействие идентификации, Command Protocols - Командные протоколы.

быстродействие или скорость идентификации (рис. 7.2). Внимание к быстродействию зависит от применения и конфигурации системы. Достаточно высокая скорость идентификации необходима в применениях, связанных с автоматизацией розничной торговли с несколькими считывателями, в то время как в применениях, связанных с проведением инвентаризации, быстродействие не столь важно. Быстродействие зависит от двух факторов: скорости передачи данных и скорости идентификации. Скорость передачи данных является собственно скоростью передачи битов данных, тогда как скорость идентификации определяется характеристиками антиколлизийного алгоритма при определении индивидуального номера метки.

7.3.1. Скорость передачи данных

Скорость передачи данных в основном зависит от периода повторения битов информации. Чем меньше период, тем больше скорость передачи данных. Следствием уменьшения периода повторения сигнала, как мы видели в главе 3, является расширение спектра сигнала. Ширина спектра ограничивается двумя факторами: требованиями частотного регламента и аппаратными ограничениями. Аппаратные ограничения, в свою очередь, определяются фильтрующими свойствами антенны и электронной схемы метки. Эти ограничения особенно заметны в системах, которые функционируют в ближней зоне поля и имеют резонансную настройку и высокую добротность. Высокая добротность приводит к сужению полосы пропускания и снижению скорости передачи данных. Антенны меток, которые работают в дальней зоне поля, также имеют конечную полосу пропускания. Как правило, однако, эти аппаратные ограничения не столь строги, как требования регламентов.

Скорость передачи данных непосредственно не связана с несущей частотой, однако регламенты выделяют на каждой частоте определенную полосу и, следовательно, этим определяется ширина спектра сигнала. Согласно требованиям регламентов более высоким частотам соответствует большая ширина полосы пропускания. Рабочая частота обычно определяет частоту синхронизации метки. В системах ближней зоны поля метки, как правило, реализуют способ синхронизации, частота которой обычно образуется делением несущей частоты. В системах дальней зоны частота синхронизации выделяется из модулирующего сигнала. Некоторые системы содержат в своем составе схемы фазовой автоподстройки частоты (ФАПЧ) и генераторы, в то время как другие системы для того, чтобы поддерживать

синхронизацию, используют фронты излучаемых сигналов. В любом случае независимо от того, каким образом формируется синхронизирующий сигнал, большей частоте синхронизации соответствует большая скорость обработки сигнала и большая потребляемая мощность. Таким образом, в системах RFID ширина спектра сигнала считывателя определяет скорость передачи данных от считывателя к метке и (в зависимости от конфигурации системы) частоту синхронизации метки, и скорость передачи данных от метки к считывателю.

Ширина спектра, обычно, задается частотным интервалом на основной несущей частоте. Этот интервал может быть задан единственным значением или несколькими значениями на различных уровнях. В зависимости от конкретной администрации или рабочей частоты регламенты определяют этот параметр различным образом. Если мы рассмотрим европейский регламент на частоте ISM 13,56 МГц, основной спектр излучения должен быть сосредоточен в частотном интервале ± 7 кГц. Побочное излучение определяется двумя уровнями. Первый – больший уровень – соответствует предельному уровню основного спектра, а второй – меньший – соответствует расстройке на ± 150 кГц. За пределами этих частот задается еще более низкий предельный уровень напряженности поля, соответствующий паразитному излучению.

В США принято простое соответствие ISM определению ± 7 кГц, а низкий уровень напряженности поля задается в виде общего фонового излучения.

В Японии приняты другие правила. Здесь полоса пропускания соответствует 7-кратной скорости передачи данных. Такой способ непосредственно связывает скорость передачи данных с требованиями электромагнитной совместимости.

Рассмотрим также ширину спектра сигналов, модулируемых меткой. Так как уровни сигнала метки значительно ниже уровня сигнала считывателя, они находятся ниже задаваемых регламентами пределов. Ширина полосы пропускания обратного канала от метки к считывателю ограничивается аппаратными возможностями. Так как считыватель является стационарным устройством, он реализует максимальные административно определенные возможности и использует различные способы кодирования и модуляции для достижения приблизительно одинаковых дальностей в различных регионах. Эти способы выбирают производители аппаратуры. В случае, если метка не способна обрабатывать сигналы многих производителей, модулированные различным образом, возникает проблема аппаратной совместимости.

В главе 3 мы видели, как кодирование и модуляция влияют на ширину спектра сигнала и скорость передачи данных. Далее, мы видели, что определенные способы кодирования позволяют получить большую энергию сигнала по сравнению с другими способами. В прямой линии от

считывателя к метке желательны сигналы, имеющие максимальную мощность, так как они переносят не только информацию, но и энергию, необходимую метке. В этом смысле скорость передачи данных и дальность связаны между собой. С точки зрения аппаратного обеспечения мы можем отметить, что для данной кодирующей и модулирующей схемы увеличение скорости передачи данных приводит к расширению спектра. С учетом ограничений по ширине спектра аппаратуры это, соответственно, приведет к снижению мощности, доступной метке. Таким образом, это дополнительно показывает взаимосвязь скорости передачи данных и дальности действия.

7.3.2. Быстродействие идентификации.

Задав скорость передачи данных, мы видим, что для определения быстродействия становятся важными антиколлизийные алгоритмы. Интегральной характеристикой алгоритмов является их эффективность, которая определяется как скорость идентификации меток поделенная на скорость передачи данных. Определив скорость передачи данных R и скорость идентификации меток T , получим эффективность алгоритма:

$$\varepsilon_p = T [\text{меток в сек.}] / R [\text{бит в сек.}] \quad (7.1)$$

Лучшие антиколлизийные схемы будут иметь более высокую эффективность. Антиколлизийные алгоритмы, конечно, влияют на затраты при проектировании аппаратуры, которые, в свою очередь, заметно определяет ее стоимость. Существует важная взаимосвязь: скорость идентификации – стоимость. Кроме того, протоколы влияют на мощность потребления – чем больше интенсивность обработки, тем больше потребляемая мощность. Как было показано в главе 6, существует три основных типа антиколлизийных алгоритмов и их несколько разновидностей. Для пассивных систем RFID низкой стоимости возможно наиболее приемлемы алгоритмы во временной области, когда метки идентифицируются детерминистическим или вероятностным методом на основе случайных задержек. Как отмечалось, каждый из методов имеет свои достоинства и недостатки. Существенным и важным преимуществом детерминистических схем является возможность разрешения коллизий и обработки сигнала за счет использования интеллектуальных ресурсов считывателя и хоста. В зависимости от применения и конфигурации могут идентифицироваться метки с определенными специфическими номерами. Детерминистические схемы способны обеспечивать селективную идентификацию. Вероятностные схемы, как правило, не имеют такого преимущества. Вернее они способны реализовать такую возможность, но только за счет дополнительного усложнения чипа метки и, следовательно, за счет увеличения его размеров и соответствующего повышения стоимости.

Необходимо также заметить, что антиколлизийные алгоритмы влияют на скорость передачи данных. Алгоритмы, требующие редкой передачи данных от считывателя, имеют меньшую полосу пропускания и, поэтому, более низкую скорость передачи данных. В этом смысле вероятностные схемы имеют преимущество, так как объем данных, передаваемых метке от считывателя, невелик.

Надежность связи в системе RFID также сильно связана с антиколлизийными алгоритмами. Надежность прямого канала связи выше, так как в нем легче обеспечить большое отношение сигнал/шум. Надежность обратного канала связи значительно ниже, поэтому более надежными являются протоколы, которые требуют передачи меньшего объема данных от метки к считывателю. Как было показано в главе 4, некоторые протоколы требуют простого распознавания метки. Однако и у таких протоколов есть недостатки – ошибка в ответе метки приводит к неправильной идентификации объекта.

Надежность каналов связи сильно связана как с дальностью действия, так и с быстродействием системы. Применение процедур обнаружения и коррекции ошибок приводит к снижению скорости передачи данных, усложняет аппаратуру, требует больших затрат мощности и, следовательно, уменьшает дальность. Как отмечалось ранее, по мере приближения к считывателю напряженность поля возрастает, сигнал становится более различимым и, следовательно, возрастает надежность связи.

Таким образом, мы видим, что и скорость передачи данных, и антиколлизийные алгоритмы влияют на быстродействие системы. Антиколлизийные протоколы также влияют на скорость передачи данных. При этом и антиколлизийные алгоритмы, и скорость передачи данных в совокупности с схемами кодирования и модуляции существенно определяют мощность потребления метки, которая, в свою очередь, влияет на дальность действия системы. И скорость передачи данных, и антиколлизийные алгоритмы особенно важны с точки зрения обеспечения надежности связи и совместимости с регламентами.

7.4. Аппаратная совместимость и стандартизация

Как отмечалось во введении, для продвижения технологии RFID на рынок важна электромагнитная совместимость регламентов различных администраций и аппаратная совместимость систем RFID различных производителей. Вопросы, касающиеся регламентов, мы рассматривали ранее. Теперь мы рассмотрим некоторые важные аспекты, относящиеся к аппаратной совместимости между системами RFID различных производителей.

В общем, с точки зрения стандартизации возможны два подхода. Системы могут быть либо полностью несовместимыми в кодировании, модуляции, командах и алгоритмах, либо полностью совместимыми.

В любом случае в настоящее время преобладает тенденция и рекомендуются системы, в которых считыватель запрашивает метку первым (системы RTF). При этом в случае принятия меткой ошибочного решения, она переходит в пассивное (молчащее) состояние и активируется только тогда, когда считыватель генерирует специальную стартовую команду. Если стартовые команды совершенно различны и низка вероятность случайной (неумышленной) генерации этой команды или сбоя в аппаратуре, можно не беспокоиться по поводу интерференции сигналов меток. Если желательно, чтобы считыватель функционировал с различными типами меток, он должен быть приспособленным для работы со всеми видами кодирования, модуляции и протоколами. Полностью совместимые системы должны иметь идентичное кодирование, модуляцию, алгоритмы и протоколы.

Достижение совместимости или, по существу, обеспечение отсутствия интерференции сигналов возможно с использованием как совместимых, так и несовместимых методов. Несовместимые методы оставляют полную свободу проектировщику, тогда как создаются трудности для конечного пользователя. Если конечный пользователь желает считывать метки RFID различных производителей, он должен устанавливать считыватели с различными схемами и протоколами. Это более дорогой путь, который имеет преимущества в обработке сигналов и позволяет избежать многих проблем совместимости. Более того, некоторые конечные пользователи предпочитают использовать метки определенных производителей. В этом смысле несовместимость позволяет избежать многих проблем. В то же время аппаратная совместимость для конечного пользователя обеспечивает недорогое и простое решение, однако для проектировщика аппаратуры оказывается наиболее сложной проблемой. При этом простор для инноваций ограничен, и конкуренция, которая обеспечивает технический прогресс и снижение стоимости аппаратуры, существенно затруднена.

Другим важным параметром, связанным с совместимостью, является рабочая несущая частота аппаратуры RFID. Очевидно, что метки, работающие на различных частотах, полностью ортогональны и вероятность интерференции их сигналов практически отсутствует. Для конечных пользователей, однако, желание считывать метки на различных частотах приводит к необходимости иметь соответствующие различные считыватели и антенны. Существует тенденция предписывать в стандартах различные несущие частоты, однако это крайне нежелательно. Поведение электромагнитных полей и волн на различных частотах и в различных применениях очень неодинаково. Конкретная частота может быть идеально

подходящей в одном применении и неприемлемой в другом, хотя и похожем применении. В то же время другая частота в таком случае окажется подходящей. Проведение опытной проверки остается важной операцией при определении поведения систем на различных частотах. Выбор единственной частоты даже для единственного применения может оказаться неудачным решением. Конкретный выбор частоты аппаратуры осуществляется на основе практической проверки применения системы RFID в реальной ситуации.

В любом смысле стандартизация в вопросах обеспечения совместимости является чувствительным фактором и может приводить к серьезному влиянию на технический прогресс технологии RFID. Достижение совместимости за счет использования метода полной несовместимости в системах RTF и полностью ортогональными стартовыми командами остается реальным подходом. Возможно также использование метода полной совместимости. Одновременно могут использоваться оба метода, оставляя право выбора производителю аппаратуры или конечному пользователю. Так как любая рабочая частота в каждом конкретном применении имеет свои собственные преимущества и недостатки, универсального решения в виде назначения какой-либо одной несущей частоты для аппаратуры RFID не существует.

7.5. Резюме

В этой главе мы рассмотрели, как важнейшие параметры системы ограничены особенностями электродинамики, связью, регламентами, аппаратным проектированием и протоколами команд. Мы увидели, что технические характеристики аппаратуры прямо или косвенно зависят от фундаментальных ограничений. Мы также рассмотрели зависимость технических характеристик и ограничений в зависимости конфигурации и от условий применения систем RFID. При проектировании меток и разработке стандартов очевидна высокая степень взаимосвязи между техническими характеристиками аппаратуры, ограничениями, конфигурацией системы и ее конкретным применением.

Глава 8

Заключение

8.1. Заключение

Последние годы характеризуются огромным потенциалом расширения различных сфер применения систем RFID, а реализация этого потенциала требует снижения их стоимости. Одновременно проводится большая работа по разработке международных и национальных стандартов технологии RFID. Уже существующие стандарты охватывают ряд применений от идентификации животных до идентификации автомобилей на автотрассах. В настоящее время прилагаются большие усилия по разработке стандартов для применения систем RFID в управлении потоками штучных товаров и в транспортной логистике. Глобальная тенденция маркирования штучных товаров низкой стоимости налагает жесткие ограничения на системы RFID. Здесь, кроме технических характеристик, важны размеры и, особенно, стоимость. Абсолютно необходима низкая стоимость меток.

Для того чтобы реализовать низкую стоимость меток, необходим специальный подход при проектировании. Для проведения грамотного проектирования крайне важно понимание основных ограничений систем RFID. Основные ограничения пассивных меток систем RFID заключаются в области электромагнетизма, в связи и в регламентах. В данной работе мы рассмотрели эти ограничения. Мы увидели, как эти ограничения влияют на размеры, стоимость и на технические характеристики – дальность действия, быстродействие, надежность связи и совместимость.

На примере меток низкой стоимости видно, что расширение возможностей и увеличение объема их выпуска приводит к снижению стоимости и продолжительности времени их использования. Усовершенствования и инновации приводят к повышению эффективности производства и снижению стоимости. Более того, технический прогресс вероятнее всего приведет к расширению функциональных возможностей: включению меток в состав различных датчиков и реализации других дополнительных функций. Кроме того, расширение возможностей приведет к разработке дополнительных применений технологии RFID, включая оперативный информационный мониторинг объектов. Повышение автоматизации процессов, в свою очередь, приведет к повышению их эффективности и продуктивности.

8.2. Направления будущих работ и исследований

При анализе основных ограничений и настоящего уровня разработок аппаратуры RFID выявилось большое число направлений будущих работ и исследований. Кратко рассмотрим некоторые из них.

Мы видели, что в пассивных системах RFID, основным фактором, ограничивающим дальность, является мощность, которой располагает метка. Хотя дальность не во всех случаях является абсолютно главным фактором, тем не менее, она важна в большом числе применений. В настоящее время источником энергии метки является электромагнитное поле, излучаемое считывателем. Хотя уже известны технологии, которые извлекают энергию из окружающей среды – «энергособирающие (energy harvesting)» технологии. Эти технологии могут устранить энергетическое ограничение и увеличить дальность, тогда ограничивающим фактором дальности станет чувствительность считывателя – приемника информационного сигнала метки. В настоящее время основная работа проводится в области реализации энергособирающих технологий.

Уже разработаны MEMS преобразователи, которые преобразуют вибрацию в электрическую энергию, показаны образцы обуви, которая извлекает энергию при ходьбе пешком и созданы интегрированные устройства, которые преобразуют рассеянную тепловую в электрическую энергию [67, 82, 83]. Метки RFID могут использовать аналогичные решения и улучшить свои параметры за счет использования окружающей рассеянной электромагнитной энергии.

Заслуживают внимания вопросы электродинамического проектирования систем RFID для специфических условий эксплуатации. В логистических и других конкретных применениях, как правило, имеется большое число окружающих предметов. Окружающие проводящие предметы могут вызывать вариацию поля и диаграмм направленности и потому существенно влиять на технические характеристики аппаратуры. Проектирование антенн допускает большое разнообразие решений, поэтому здесь конечно необходимо проведение исследований.

Несмотря на то, что к настоящему времени достигнут существенный прогресс в области разработки и оптимизации канала передачи данных, однако в части кодирования и модуляции сигнала остается еще большое поле деятельности. В различных странах имеются отличия в ограничениях ширины спектра и напряженности поля в различных диапазонах несущих частот. Для оптимизации систем RFID необходимо проведение работ по международной гармонизации спектра, а также по исследованию вопросов кодирования, модуляции и фильтрации сигналов. Кроме того, необходим прогресс в области проектирования недорогих FSK и PSK детекторов; также будут полезны улучшения и в генераторной схемотехнике. Целесообразно

проведение разработок новых антиколлизийных алгоритмов в пространственной и частотной областях.

Необходимо проведение исследований в области использования новейших достижений в антенной технике. Многообещающим может быть использование фрактальных антенн, так как это может обеспечить уменьшение габаритов меток в сочетании с высокими техническими характеристиками [84]. Должны быть проведены исследования, касающиеся использования объектов со встроенными метками. Упаковка некоторых объектов может использоваться в качестве интегрированной антенны метки или в качестве накопителя энергии. Также должны быть разработаны новые производственные технологии, которые позволят обеспечить изготовление и сборку интегрированных меток, которые состоят из антенны и электронного чипа.

Должны быть исследованы возможности реализации бесчиповых меток, созданию которых в настоящее время уделяется большое внимание. Особый интерес вызывают исследования, касающиеся дальности, памяти и антиколлизийных возможностям таких меток.

Хотя основное внимание данной работы сосредоточено на метке, также большой объем исследований необходимо провести по считывателям. Особенно это касается рассмотрения коллизийных ситуаций между считывателями. Из-за увеличения количества плотность считывателей RFID возрастает и, соответственно, возрастает вероятность их взаимовлияния друг на друга. Такое помеховое взаимное влияние между считывателями особенно вредно при необходимости обеспечения высокой надежности связи между метками и считывателями.

Таким образом новые применения и дополнительные возможности технологии RFID способны стимулировать большое число интересных и практически необходимых исследований. В большом числе современных и будущих применений очень полезной может быть интеграция информационных датчиков, к которым относится аппаратура RFID, с устройствами использования различных видов энергии, или при помощи MEMS устройств, или при помощи других, более простых материальных структур. Могут исследоваться новые применения, связанные с робототехникой и автоматизацией. При этом автоматическая идентификация радиочастотных меток в сочетании с сетевыми базами данных обеспечит оперативное распознавание объектов информации и, соответственно, управление производственными технологическими процессами.

Список литературы

- [1] Yogesh Joshi, "Information Visibility and Its Effect on Supply Chain Dynamics," M.S. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 2000, <<http://auto-id.mit.edu/research/whitepapers.html>>.
- [2] MIT Auto-ID Center, "The Networked Physical World: Proposals for engineering the next generation of computing, commerce, and automatic-identification," MIT-AUTOID-WH-001, 2000, <<http://auto-id.mit.edu/research/whitepapers.html>>
- [3] Jim Eagle, "RFID: The Early Years 1980-1990," 2001, <<http://members.surfbest.net/eaglesnest/rfidhist.htm>>
- [4] Alfred R. Koelle., Steven W. Depp, Jeremy A. Landt, and Ronald E. Bobbett, "Short-Range Passive Telemetry by Modulated Backscatter of Incident CW RF Carrier Beam," Biotelemetry, Vol. 3., Academic Press, New York, pp. 337-340, 1976.
- [5] Thierry Roz and Vincent Fuentes, "Using Low Power Transponders and Tags for RFID Applications," EM Microelectronic Marin SA, Marin, Switzerland, <<http://www.emmarin.ch/>>.
- [6] Klaus Finkenzeller, RFID Handbook, John Wiley and Sons, Inc., New York, 1999.
- [7] John R. Tuttle, "Traditional and Emerging Technologies and Applications in the Radio Frequency Identification (RFID) Industry," IEEE Radio Frequency Integrated Circuits Symposium, 1997.
- [8] S. Shuji, H. Yuzo, A. Dobashi, M. Okumura, and T. Kusuzaki, "Products Lifecycle Management System Using Radio Frequency Identification Technology," Proceedings of the 7th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vol. 2, 1999, pp. 1459-1467.
- [9] David L. Brock, "The Electronic Product Code," MIT-AUTOID-WH-002, 2001, <<http://auto-id.mit.edu/research/whitepapers.html>>.
- [10] Peter J. Harrop, Chipless Smart Labels: The Ultimate RFID, IDTechx, Cambridge, UK, 1988.
- [11] N. Komatsu, T. Kon, H. Mukai, K. Kamigama, "Research and Study of Elementary Technology for Achieving Greater Efficiency in Postal Processing: The Present State of RFID and their Applications to Postal Services," IPTP Monthly Report, No. 119, 1998.

- [12] David H. Staelin, Ann. W. Morgenthaler, and Jin Au Kong, *Electromagnetic Waves*, Prentice Hall, New Jersey, 1994.
- [13] Warren L. Stutzman and Gary A. Thiele, *Antenna Theory and Design*, 2nd ed., John Wiley and Sons, Inc., New York, 1998.
- [14] David K. Cheng, *Field and Wave Electromagnetics*, 2nd ed., Addison-Wesley Publishing Co., New York, 1992.
- [15] Richard Ribon Fletcher, "A Low-Cost Electromagnetic Tagging Technology for Wireless Identification, Sensing, and Tracking of Objects," M.S. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1997.
- [16] IEEE Standard Dictionary of Electrical and Electronics Terms, 3rd ed., ANSI/IEEE Standard 100-1984.
- [17] Eugene F. Knott, John F. Shaeffer, and Michael T. Tuley, *Radar Cross Section*, 2nd ed., Artech House, Boston, 1993.
- [18] R. C. Hansen, "Relationships Between Antennas as Scatterers and as Radiators," *Proceedings of the IEEE*, Vol. 77, No. 5, pp. 659-652, 1989.
- [19] Robert Blair Green, "The General Theory of Antenna Scattering," Ph.D. Thesis, The Ohio State University, Columbus, OH, 1963.
- [20] Alan Demers, Scott Elrod, Christopher Kantarjiev, and Edward Richley, "A Nano-Cellular Local Area Network Using Near-Field RF Coupling," Xerox Corporation, Palo Alto, 1994.
- [21] Theodore S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, New Jersey, 1996.
- [22] Jorgen Bach Andersen, Theodore S. Rappaport, and Susumu Yoshida, 1995, "Propagation Measurements and Models for Wireless Communications Channels," *IEEE Communications Magazine*, January, 1995.
- [23] John B. Anderson, *Digital Transmission Engineering*, IEEE Press, New York, 1999.
- [24] Leon W. Couch, *Digital and Analog Communications Systems*, 6th ed., Prentice-Hall, New Jersey, 2001.
- [25] Joseph L. LoCicero and Bhasker P. Patel, "Line Coding," *The Mobile Communications Handbook*, Gibson, Jerry D., ed., CRC Press, Boca Raton, Florida, 1999, pp. 6-1 - 6-18.

- [26] Hwei P. Hsu, *Probability, Random Variables and Random Processes*, McGraw-Hill, New York, 1997.
- [27] William C. Lindsey, and Marvin K Simon, *Telecommunication Systems Engineering*, Prentice-Hall, New Jersey, 1973.
- [28] R.C. Titsworth and L. R. Welch, "Power Spectra of Signals Modulated by Random and Pseudorandom Sequences," Jet Propulsion Laboratory, Pasadena, CA, Technical Report No. 32-140, October 1961.
- [29] H. Hecht and A. Guida, "Delay Modulation" *Proceedings of the IEEE*, Vol. 57, No. 7, July 1969, pp.1314-1316.
- [30] E. N. Gilbert, "Capacity of a Burst-Noise Channel," *Bell System Technical Journal*, Vol. 39, pp. 1253-1266, September 1960.
- [31] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Error Channels," *Bell System Technical Journal*, Vol. 42, p. 1997, September 1963.
- [32] Michele Zorzi and Ramesh R. Rao, "Error Control Strategies for the Wireless Channel," *IEEE 5th International Conference on Universal Personal Communications*, Vol. 2, 1996, pp. 970-974.
- [33] J. R. Yee and E. J. Weldon, "Evaluation of the Performance of Error-Correcting Codes on a Gilbert Channel," *IEEE Transactions on Communications*, Vol. COM-43, August 1995, pp. 2316-2323.
- [34] Mischa Schwartz, *Information, Transmission, Modulation, and Noise*, McGraw-Hill, New York, 1980.
- [35] David Withers, *Radio Spectrum Management: Management of the Spectrum and Regulation of Radio Services*, 2nd ed., The Institution of Electrical Engineers, London, 1999.
- [36] ITU RR, 1998, *Radio Regulations*, Vol. 1., International Telecommunications Union.
- [37] ERO, European Radiocommunications Office, <<http://www.ero.dk>>.
- [38] ETSI, European Telecommunications Standards Institute, <<http://www.etsi.org>>.
- [39] FCC, Federal Communications Commission, <<http://www.fcc.gov>>.
- [40] MPT, Ministry of Post and Telecommunications, <<http://www.joho.soumu.go.jp>>.

- [41] ARIB, Association of Radio Industries and Businesses, <<http://www.arib.or.jp>>.
- [42] ITU-R 213/1, Technical and Operating Parameters and Spectrum Requirements for Short-Range Radiocommunication Devices, 2001.
- [43] ITU-R SM.1056, Limitation of Radiation for Industrial, Scientific and Medical (ISM) Equipment, 1994.
- [44] CEPT 70-03, Relating to the Use of Short Range Devices (SRD), May 2001.
- [45] ETSI EN 300330-1 V1.3.1 (2001-04), Electromagnetic Compatibility and Radio Spectrum Matters (ERM); Short Range Devices (SRD); Radio equipment in the frequency range 9 kHz to 25 MHz and inductive loop systems in the frequency range 9 kHz to 30 MHz; Part 1: Technical characteristics and test methods, 2001.
- [46] FCC Part 15, Radio Frequency Devices, 2001.
- [47] David Mawdsley, EMC Conducted Emissions Measurement: Detectors... what, why and when?," Technical Notes from Laplace Instruments, 2001, <http://perso.wanadoo.fr/dicomtech/article_lap_2.htm>.
- [48] CISPR 16-1, Specification for Radio Disturbance and Immunity Measuring Apparatus and Methods - Part 1: Radio Disturbance and Immunity Measuring Apparatus, International Electrotechnical Commission. 1999.
- [49] ETSI EN 300220-1 V1.3.1 (2000-09), Electromagnetic Compatibility and Radio Spectrum Matters (ERM); Short Range Devices (SRD); Radio equipment to be used in the 25 MHz to 1000 MHz frequency range with power levels ranging up to 500 mW; Part 1: Technical characteristics and test methods, 2000.
- [50] ETSI EN 300440-1 V1.3.1 (2000-07), Electromagnetic Compatibility and Radio Spectrum Matters (ERM); Short Range Devices; Radio equipment to be used in the 1 GHz to 40 GHz frequency range; Part 1: Technical characteristics and test methods, 2000.
- [51] John Burns, Richard Rudd, and Zoran Spasojevic, "Compatibility Between Radiocommunication and ISM Systems in the 2,4 GHz Frequency Band," Aegis Systems Limited, 1999.
- [52] Acheson Colloids Company, "Product Selection Guide," <<http://www.achesoncolloids.com/>>.
- [53] Ramesh Garg, Prakash Bhartia, Inder Bahl, Apisak Ittipiboon, Microstrip Antenna Design Handbook, Artech House, Boston, 2000.

- [54] D. Friedman, H. Heinrich, D-W. Duan., "A Low-Power CMOS Integrated Circuit for Field-Powered Radio Frequency Identification Tags," IEEE International Solid-State Circuits Conference, 1997, p. 294.
- [55] Ulrich Kaiser and Wolfgang Steinhagen, "A Low-Power Transponder IC for High-Performance Identification Systems," IEEE Journal of Solid-State Circuits, Vol. 30, No. 3, pp. 306–310, March 1995.
- [56] Stephane Donnay, Philip Pieters, Kristof Vaesen, Wim Diels, Piet Wambacq, Walter De Raedt, Eric Beyne, Marc Engels, and Ivo Bolsens, "Chip-Package Codesign of a, Low-Power 5-GHz RF Front End," Proceedings of the IEEE, Vol. 88, No. 10, pp. 1583–1597, October 2000.
- [57] M.H. Brodsky, Progress in Gallium Arsenide Semiconductors. Scientific American, pp.68–72, February 1990.
- [58] J.D. Meindl, Chips for Advanced Computing. Scientific American, pp. 78-88, October 1987.
- [59] Lawrence E. Larson, "Integrated Circuit Technology Options for RFID's - Present Status and Future Directions," IEEE Custom Integrated Circuits Conference, 1997, pp. 169–176.
- [60] Thomas H. Lee, The Design of CMOS Radio-Frequency Integrated Circuits, Cambridge University Press, Cambridge, UK, 1998.
- [61] Agilent AN 1089, "Designing Detectors for RFID Tags," Agilent Application Note, 1999.
- [62] Paul Horowitz and Winfield Hill, The Art of Electronics, 2nd ed., Cambridge University Press, Cambridge, UK, 1989.
- [63] Eric A. Vittoz, "Micropower IC," Low-Power CMOS Design, Anantha Chandrakasan, Robert Brodersen, ed, IEEE Press, New York, 1998.
- [64] Chenming Hu, "Device and Technology Impact on Low Power Electronics," Low Power Design Methodologies, Kluwer Academic Publishers, Boston, 1996, pp. 21–33.
- [65] Jan M. Rabaey, Massoud Pedram and Paul Landman, "Introduction," Low Power Design Methodologies, Kluwer Academic Publishers, Boston, 1996, pp. 1–15.
- [66] P. Vanoostende, et al., "Evaluation of the limitations of the simple CMOS power estimation formula: comparison with accurate estimation," Proceedings of the European Workshop on Power and Timing Modeling, 1992, pp. 16–25.

- [67] Athas, William C., "Energy-Recovery CMOS," Low Power Design Methodologies, Kluwer Academic Publishers, Boston, 1996.
- [68] Benny Bing, Broadband Wireless Access, Boston, Kluwer Academic Publishers, MA, 2000.
- [69] Peter Hawkes, "Anti-collision and Transponder Selection Methods for Grouped "Vicinity" Cards and RFID tags," IEE Colloquium on RFID Technology, 1999, pp. 7/1–7/12.
- [70] Josef Schuermann, U.S. Patent: System and method for reading multiple RF-ID transponders, Patent Number: US5500651, March 1996.
- [71] Robert Fleming, Cherie Kushner, "Integrated, Low-Power, Ultra-Wideband Transceivers for Distributed Position Location and Communication," Aether Wire & Location, Inc., January 1998, <http://www.aetherwire.com/Aether_Wire/Technical_Abstracts_97.pdf>.
- [72] Konstantin O. Aslanidis, Alain Berthon, U.S. Patent: Method for repeating interrogations until failing to receive unintelligible responses to identify plurality of transponders by an interrogator, Patent Number: US5929801, July 1999.
- [73] Ching Law, Kayi Lee, Kai-Yeung Siu, "Efficient Memoryless Protocol for Tag Identification," In Proceedings of the 4th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications, August 2000.
- [74] Marcel Jacomet, Adrian Ehram, Urs Gehrig, "Contactless Identification Device With Anticollision Algorithm," IEEE Computer Society, CSCC'99, Conference on Circuits, Systems, Computers and Communications, 4–8 July 1999 Athens, <http://www.microlab.ch/academics/r_and_d/research/csc/csc99.pdf>.
- [75] Ronald Walter and Keith Vertrees, U.S. Patent: Method and apparatus for radiofrequency identification tags, Patent Number: US5856788, January 1999.
- [76] Shigeru Furuta, U.S. Patent: Method and system for identifying and communicating with a plurality of contactless IC cards, Patent Number: US5698837, December 1997.
- [77] EM Microelectronic-Marin SA, P4022: Multi Frequency Contactless Identification Device, <http://www.emmarin.com/pdf/RFID/P4022_C.pdf>.
- [78] Microchip Technology Inc., microID 125 kHz Design Guide, <<http://www.microchip.com/Download/lit/pline/rfid/guides/51115e.pdf>>.

[79] Donald L. Black. Dale Yornes, U.S. Patent: Method for resolving signals collisions between multiple RFID transponders in a field, Patent Number: US6091319, July 2000.

[80] Pourang Mahdavi, European Patent: A system and method for communicating with multiple transponders, Patent Number: EP1017005A2, May 2000.

[81] John H. Bowers, John Nelson Daigle, Rade Petrovic, U.S. Patent: Anticollision protocol for reading multiple RFID tags, Patent Number: US5883582, March 1999.

[82] Scott Meninger, Jose Oscar Mur-Miranda, Rajeevan Amirtharajah, Anantha Chandrakasan and Jeffrey Lang, "Vibration-to-Electric Energy Conversion," 1999 International Symposium on Low-Power Electronic Design (ISLPED '99), July 1999, pp. 48–53.

[83] N. S. Shenck and J. A. Paradiso, "Energy Scavenging with Shoe-Mounted Piezoelectrics," IEEE Micro, Vol. 21, No. 3, May-June 2001.

[84] R. Breden and R. J. Langley, "Printed Fractal Antennas," IEE National Conference on Antennas and Propagation, 1999.

Приложение 1

С.В. Корнеев

К вопросу о дальности действия систем RFID

Одним из основных параметров систем RFID является дальность действия. В индуктивных или емкостных системах, функционирующих в ближней зоне поля, дальность действия невысока – она достигает единиц – десятков сантиметров. Причиной тому является резкое спадание напряженности электрического (обратная кубическая зависимость) и магнитного (обратная квадратичная зависимость) полей с расстоянием в совокупности с необходимостью соблюдения строгих требований по электромагнитной совместимости (регламентов) и санитарных норм. При этом возможности увеличения дальности таких систем практически уже исчерпаны.

Дальность действия систем RFID, функционирующих в дальней зоне электромагнитного поля в зависимости от типа метки – чиповая или бесчиповая, пассивная или полупассивная – может составлять от единиц до десятков метров. По сравнению с системами ближней зоны дальность увеличивается за счет менее резкого (линейного) уменьшения напряженности электромагнитного поля.

Кроме положительного эффекта, увеличение дальности действия может приводить к нежелательным последствиям. Так, в частности, увеличение дальности действия приводит к расширению зоны считывания аппаратуры и может снижать ее быстродействие из-за необходимости разрешения коллизий сигналов от большого числа меток. Кроме того, могут появиться проблемы электромагнитной совместимости с другим радиоэлектронным оборудованием, например, с близко расположенными считывателями других систем RFID. Следовательно грамотный выбор параметров, влияющих на дальность действия, может повысить эффективность функционирования систем RFID в целом.

1. Энергетическое и модуляционное уравнения дальности

Рассмотрим зависимость дальности действия наиболее широко распространенного в мире типа аппаратуры – с совмещенной приемной и передающей антенной считывателя и пассивными чиповыми метками от характеристик и параметров этой аппаратуры. Дальность действия таких систем описывается двумя уравнениями [1].

Первое из них, так называемое «энергетическое» уравнение, характеризует прямую линию (forward line) – канал доставки энергии от считывателя к метке, необходимой для обеспечения энергопитания активных элементов чипа.

Именно это уравнение, как правило, ограничивает дальность действия систем RFID с чиповыми метками:

$$R_E = (k_E \cdot P \cdot G_R \cdot A_e / 4\pi S)^{1/2}, \quad (1)$$

где R_E – энергетическая дальность считывания метки,
 P – мощность генератора считывателя,
 G_R – коэффициент усиления антенны считывателя,
 A_e и S – эффективная площадь антенны и чувствительность метки соответственно,
 k_E – энергетический коэффициент.

В отличие от известного уравнения радиосвязи в формулу (1) введен энергетический коэффициент k_E . Этот коэффициент учитывает соотношение между мощностью сигнала, поступающей в нагрузку антенны (чип), и мощностью, которая по принципу действия технологии RFID отражается меткой и несет информацию об идентифицируемом объекте. Чувствительность S является мощностью на выходе антенны метки, которая гарантирует минимально необходимое энергопитание подсоединенного к антенне чипа.

Второе – «модуляционное» уравнение, описывает обратную линию (reverse line), т.е. прием считывателем отраженного от метки сигнала, модулированного кодами идентификационного номера или данных метки:

$$R_M = [P \cdot G_R^2 \cdot \lambda^2 \cdot \Delta\sigma / (4\pi)^3 \cdot S_R]^{1/4}, \quad (2)$$

где R_M – модуляционная дальность считывания метки,
 λ – длина электромагнитной волны сигнала,
 $\Delta\sigma = (\sigma_{max} - \sigma_{min})$ – вариация ЭПР антенны метки,
 S_R – рабочая чувствительность приемника считывателя.

Параметры σ_{max} и σ_{min} являются максимальным и минимальным значениями ЭПР, которые обеспечиваются вариацией нагрузки (чипом) и модуляцией обратного рассеяния антенны метки. Специфика уравнения (2) состоит в том, что в технологии RFID оперируют с вариацией ЭПР. Исключение процесса модуляции ЭПР (т.е. $\Delta\sigma = 0$ и $\sigma_{max} = \sigma$) приводит это уравнение к традиционному виду уравнения радиолокации. Рабочая чувствительность считывателя S_R учитывает соотношение сигнал/шум, необходимое для достижения заданной достоверности считывания (безошибочной идентификации) номера метки. Расчеты показывают, что для 100-битного

номера и достоверности считывания метки 0,9999 соотношение сигнал/шум должно быть не менее 20 дБ.

Энергетическое уравнение (1) характеризует максимальную дальность действия аппаратуры RFID с пассивными чиповыми метками, которая достигается при угловом и поляризационном согласовании антенн считывателя и метки, а также при согласовании импедансов антенны и чипа метки. Модуляционное уравнение (2) характеризует максимальную дальность действия аппаратуры RFID с полуактивными метками, у которых энергопотребление чипа обеспечивается элементом питания.

2. Зависимость дальности действия от несущей частоты

В настоящее время технологии RFID в UHF и микроволновом диапазоне выделены радиочастоты 433 МГц, 860 – 960 МГц, 2,45. ГГц и 5,78 ГГц. Анализ технических характеристик аппаратуры различных производителей показывает, что дальность ее действия существенно различается в зависимости от диапазона несущих частот. С целью лучшей ориентации системных интеграторов и конечных пользователей в процессе выбора конкретной аппаратуры целесообразно рассмотреть зависимость дальности действия систем RFID от несущей частоты или длины волны сигнала. Для этого примем два предположения:

– Параметры P и G_R считывателя будем считать фиксированными. Данное предположение позволяет сравнивать системы с равными энергетическими потенциалами и пространственными зонами считывания.

– Эффективная площадь A_e и вариация ЭПР $\Delta\sigma$ антенны метки пропорциональны квадрату длины волны сигнала λ , т.е. $A_e \approx k_1 \cdot \lambda^2$ и $\Delta\sigma \approx k_2 \cdot \lambda^2$, где k_1 и k_2 – коэффициенты пропорциональности. Данное предположение справедливо в большинстве практических случаев [2, 3].

С учетом этого уравнения (1) и (2) можно записать в следующем виде:

$$R_E \approx \lambda (k_E \cdot P \cdot G_R \cdot k_1 / 4\pi S)^{1/2} \sim \lambda,$$

$$R_M \approx \lambda [P \cdot G_R^2 \cdot k_2 / (4\pi)^3 \cdot S_R]^{1/4} \sim \lambda.$$

Следовательно, при прочих равных условиях дальность действия систем радиочастотной идентификации с пассивными и полуактивными чиповыми метками приблизительно пропорциональна длине волны сигнала. Такой вывод соответствует данным из технических характеристик аппаратуры различных производителей: в диапазоне 900 МГц дальность действия составляет 5...6 м, а в диапазоне 2,45 ГГц – 1,5...2 м, при этом соотношение длин волн этих диапазонов близко к 3. С точки зрения достижения максимальной дальности предпочтителен диапазон 433 МГц. Однако здесь

для обеспечения эффективности излучения требуются слишком большие габариты антенн. Кроме того, в связи с малой шириной выделенного диапазона частот могут возникать проблемы электромагнитной совместимости. По этим причинам в настоящее время наибольшее распространение находят системы RFID UHF диапазона 860 – 960 МГц.

3. Зависимость дальности действия от глубины модуляции

Глубина модуляции сигнала в системах RFID не выступает в качестве основной технической характеристики аппаратуры. Этот параметр фигурирует лишь в описаниях радиоинтерфейсов – протоколов обмена командами и данными между считывателем и меткой. Однако глубина модуляции сигнала определяет величину энергетического коэффициента k_E , который, согласно уравнению (1), непосредственно влияет на энергетическую дальность действия R_E системы. От глубины модуляции зависит и $\Delta\sigma$ – вариация ЭПР антенны метки, что в соответствии с формулой (2) влияет на модуляционную дальность действия R_M аппаратуры. Кроме того, глубина модуляции определяет ширину спектра сигнала, который является важным параметром при обеспечении административных норм электромагнитной совместимости (ЭМС). Проведем отдельный анализ влияния глубины амплитудной модуляции сигнала в прямой линии M_f и в обратной линии M_r на дальность действия аппаратуры RFID. Для проведения количественных оценок будем использовать международный стандарт ISO 18000-6, в котором предписываются номинальные значения глубины модуляции сигнала в прямой линии для варианта А – 30% и для варианта В – 15% или 99%.

В прямой линии M_f является параметром сигнала считывателя. Его величина может изменяться в пределах от 0 до 100%. Можно показать [1], что коэффициент k_E связан с глубиной модуляции зависимостью:

$$k_E = [(1 - M_f)^2 + Q \cdot M_f^2], \quad (3)$$

где Q – скважность кодового сигнала.

На рис.1. показана зависимость коэффициента k_E от глубины модуляции M_f для различных значений Q .

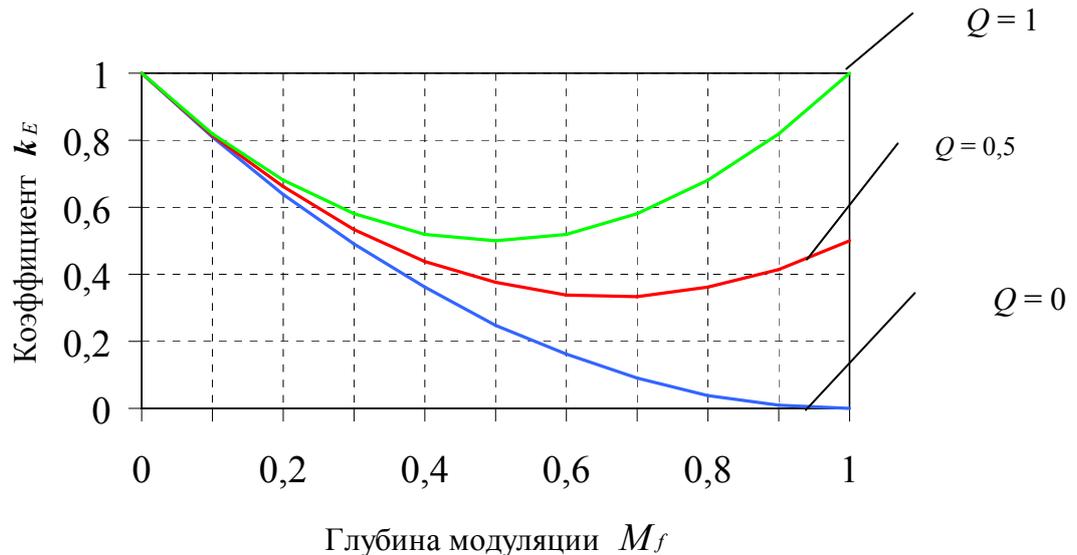


Рис.1. Зависимость коэффициента k_E от глубины модуляции M_f .

Согласно ISO18000 выбор конкретного значения скважности модулирующего сигнала остается за национальными администрациями, однако практически удобно использовать кодирование с $Q \approx 0,5$. При произвольной глубине модуляции и $Q = 0,5$ энергетический коэффициент k_E равен:

$$k_E = 1 - 2 M_f + 1,5 M_f^2. \quad (4)$$

Из рассмотрения рис.1 видно, что в этом случае k_E минимален при $M_f = 0,65$ и согласно уравнению (1) применение модуляции такой глубины с точки зрения энергообеспечения метки наименее выгодно. При величине $M_f = 1$, соответствующей модуляции ООК, которая реализуется в большинстве систем RFID, $k_E = 0,5$.

Теперь обратимся к уравнению модуляционной дальности (2). Можно показать [1], что вариация ЭПР $\Delta\sigma = \sigma_{max} \cdot M_r^2$, где M_r – глубина модуляции сигнала на входе антенны метки. Тогда дальность считывания пассивных меток в в обратной линии системы RFID примет следующий вид:

$$R_M = [P \cdot G_R^2 \cdot \lambda^2 \cdot \sigma_{max} \cdot M_r^2 / (4\pi)^3 \cdot S_R]^{1/4}. \quad (5)$$

Также можно показать [5], что при манипуляции активными импедансами в нагрузке полуволнового вибратора $\Delta\sigma < 0,75 \sigma_{max}$, т.е. максимально достижимая глубина амплитудной модуляции $M_{r\ max} = (0,75)^{0,5} \approx 0,866$.

В результате расчета по формулам (1) и (2) для типичных значений параметров стационарного считывателя в диапазоне 900 МГц ($P = 1$ Вт, $G_R = 4$, $\lambda = 30$ см, $S_R = -100$ дБВт, $k_E = 0,5$) и меток с антеннами в виде

полуволновых диполей ($A_e = 0,01 \text{ м}^2$, $\Delta\sigma = 0,06 \text{ м}^2$, $S = 30 \text{ мкВт}$) позволяет получить: энергетическая дальность $R_E \approx 7$ метров, а модуляционная дальность $R_M \approx 25$ метров. Таким образом для грамотного проектирования аппаратуры RFID необходимо оперировать двумя параметрами – энергетической и модуляционной дальностью. При этом дальность действия систем с полуактивными метками существенно больше дальности действия систем с пассивными метками. Повышение дальности, однако, достигается за счет усложнения, удорожания и сокращения срока службы метки. Именно поэтому пассивные метки в настоящее время находят значительно более широкое применение.

4. Определение технических параметров аппаратуры

Использование уравнений (1) – (5) позволяет в процессе проектирования определять технические параметры аппаратуры RFID, которые обеспечивают необходимую дальность действия.

Первым необходимым условием достижения заданной дальности R действия аппаратуры RFID с пассивными чиповыми метками является выбор параметров, соответствующих энергетическому уравнению (1). Проектирование считывателя, в результате которого определяются параметры λ , P , G_R и S_R , мало отличается от проектирования обычных приемных и передающих устройств. Из-за технических и стоимостных ограничений наиболее сложной задачей является проектирование метки (т.е. определение параметров A_e , S и $\Delta\sigma$). Поскольку S – минимальная рабочая чувствительность чипа, определена его производителями, а число производителей в мире чипов UHF диапазона крайне невелико, в первую очередь проектирование метки сводится к проектированию ее антенны с необходимой эффективной площадью A_e и вариацией ЭПР $\Delta\sigma$.

На основании уравнения (1) первое условие обеспечения $R \geq R_E$ можно записать следующим образом:

$$A_e \geq 4\pi R^2 S / k_E P G_R. \quad (6)$$

Вообще безразмерная величина ($S/P \cdot G_R$) может являться интегральным параметром, который характеризует эффективность аппаратуры в целом и по которому можно сравнивать системы RFID с пассивными метками различных производителей.

Дальность действия систем RFID с пассивными метками, однако, зависит не только от обеспечения энергопитания чипа. При снижении чувствительности приемника S_R или вариации ЭПР метки $\Delta\sigma$ возникает ситуация, когда $R_M \leq R_E$. В этом случае дальность действия системы с пассивной меткой будет определяться модуляционным уравнением (2).

Тогда, приравняв правые части уравнений (1) и (2), для систем, у которых антенна метки является полуволновым диполем, получим второе условие обеспечения заданной дальности $R \geq R_M = R_E$:

$$\Delta \sigma \geq k_E^2 \cdot P \cdot S_R A_e G_D / S^2, \quad (7)$$

где G_D – усиление полуволнового диполя.

В результате, исходя из заданной дальности действия R и известных параметров аппаратуры (λ , P , G_R , S_R и S), можно определить два основных параметра метки: A_e и $\Delta \sigma$ – эффективную площадь и вариацию ЭПР ее антенны соответственно.

Таким образом, использование двух уравнений радиочастотной идентификации позволяет обеспечить проектирование систем RFID с пассивными и полупассивными метками для различных условий применения.

Список литературы

1. Корнеев С.В. "Оптимизация параметров технологии радиочастотной идентификации", Радиопромышленность, № 3, 2002, С. 39-44.
2. Кобак В.О. Радиолокационные отражатели. М., "Советское радио", 1975.
3. Марков Г.Т. Антенны. М.-Л., Госэнергоиздат, 1960.
4. Reynolds M. "Microwave RFID: Passive Scattering and Active Transponders", MIT, 2002.
5. Корнеев С.В., Рунге А.В. "К вопросу об управлении эффективной поверхностью рассеяния диполей в технологии радиочастотной идентификации". В кн.: Антенны / Под ред. Л.Д. Бахраха. Вып. 6. - М.: Радио и связь, 2002. - С.56-60.

Приложение 2

Системы RFID с использованием нейронных сетей

До настоящего времени проблема разрешения коллизий сигналов меток, которые одновременно находятся в поле считывания (коллизионный арбитраж), остается чрезвычайно актуальной. Практически все известные антиколлизионные алгоритмы реализованы во временной области. Независимо от используемых методов и протоколов (RTF или TTF, бинарные или вероятностные, определенные наиболее современным протоколом в международном стандарте ISO 18000-6) такие алгоритмы приводят к усложнению и, следовательно, удорожанию метки, а также к снижению быстродействия и дальности действия систем RFID с пассивными метками. Более того, в этом случае абсолютно неразрешимой является задача коллизионного арбитража сигналов в системах с бесчиповыми метками, изготовленными, в частности, с использованием технологии поверхностных акустических волн (ПАВ). В то же время бесчиповые метки перспективны из-за потенциально низкой стоимости и возможности обеспечения повышенной дальности действия системы. Кроме того, ПАВ метки обладают несомненным преимуществом в части высокой устойчивости к эксплуатационным воздействиям (температура, радиация, электромагнитные поля и т.д.).

В этой связи чрезвычайно актуально появление перспективы решения проблемы коллизионного арбитража с использованием нейросетевых алгоритмов. Для иллюстрации плодотворности такого подхода ниже приводится перевод статьи авторов Y. Deville, J. Damour, N. Charkani. "Multi-tag radio-frequency identification systems based on new blind source separation neural networks". *Neurocomputing*, #49 (2002), 369–388. Научное редактирование перевода выполнено д.т.н. проф. А.И. Галушкиным.

Y. Deville, J. Damour, N. Charkani.

Системы радиочастотной идентификации множества меток на основе нового слепого метода разделения источников с использованием нейронных сетей

Аннотация

В повседневных применениях при идентификации объектов или людей электронные системы интенсивно вытесняют механические устройства или ручные операции.

© Y. Deville, J. Damour, N. Charkani, 2002

©С.В. Корнеев, перевод на русский язык, 2005

В настоящее время наиболее предпочтительными являются системы бесконтактной радиочастотной идентификации, однако в этом случае имеются трудности при одновременной обработке сигналов нескольких объектов. В настоящей работе описывается решение этой проблемы, основанное на технологии слепого разделения источников. Эффективность такого подхода продемонстрирована экспериментально с использованием аппаратуры, реализующей предложенную систему с применением цифровой обработки сигнала в реальном масштабе времени. Проведено сравнение различных способов разделения при помощи нейронных сетей, а также показано, что наиболее привлекательны недавно предложенные решения, благодаря их простоте, хорошим характеристикам и самонормирующемуся (т.е. автоматическому) функционированию.

1. Введение

Во многих реальных ситуациях необходимо идентифицировать людей, животных или иные объекты. Типичными примерами являются идентификация владельца автомобиля перед запуском его двигателя, ограничение доступа в определенные территории, идентификация рогатого скота или контроль потока производимой продукции на предприятиях. В прошлом для реализации таких задач использовались подходы, в основном с применением механических устройств (таких как ключи для запуска автомобильных двигателей) или операций с использованием людей (т.е. визуальный контроль людей, рогатого скота или продукции в приведенных выше примерах). Такие подходы интенсивно вытесняются различными разновидностями электронных устройств, и особенно системами, основанными на использовании технологии радиочастотной идентификации (RFID).

Пример такой системы RFID [10, 19, 20, 22] показан на рис.1. Она состоит из базовой станции (считывателя) и небольших идентифицирующих устройств (меток), которые содержат резонансный LC - контур, контроллер и программируемую память (EEPROM). Содержимое памяти специфично для каждой метки и позволяет идентифицировать носителя метки (человека или объект).

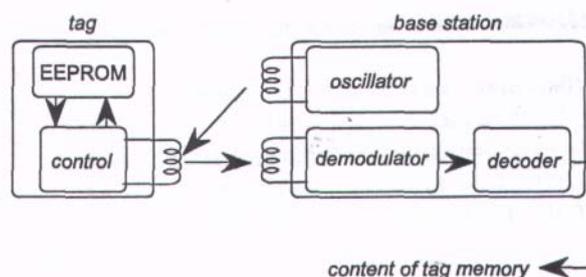


Рис.1. Система радиочастотной идентификации с одной меткой.
 tag – метка, base station – считыватель, oscillator – генератор, control – управление, demodulator – демодулятор, decoder – декодер, content of tag memory – содержимое памяти метки.

Основной принцип работы такой системы сводится к следующему. Считыватель излучает радиоволну, которая принимается единственной меткой. Метка, таким образом, получает энергию и отражает радиоволну той же частоты (благодаря индуктивной связи), модулированную кодированным содержимым памяти. Считыватель принимает этот сигнал, демодулирует и декодирует его, чтобы определить содержимое памяти (более подробно процедура кодирования рассмотрена в Приложении В). Затем идентификационная система верхнего уровня проверяет эти данные и, соответственно, управляет процессом.

Привлекательность такой системы состоит в том, что она обеспечивает бесконтактное взаимодействие между считывателем и метками (избегая, таким образом, ограничений на позиционирование объекта с меткой), причем метки не требуют источника питания. Однако, когда в поле считывателя находятся две метки, они обе отвечают на излученный считывателем сигнал. При этом демодулированный сигнал считывателя является смесью двух компонент от двух меток и не может быть декодирован. Такая система неспособна одновременно идентифицировать два объекта. Известны несколько способов решения этой проблемы. Некоторые из них состоят в том, что считыватель и метки взаимодействуют в соответствии с заранее определенным протоколом, так что сигналы каждой метки успешно разделяются [4]. Такой подход снижает быстродействие и усложняет систему, так как для реализации протокола необходимы существенные дополнительные аппаратные и программные средства и в считывателе, и в метках. Поэтому такое решение не привлекательно. Другой подход состоит в использовании меток на различных частотах [12]. Это вновь приводит к усложнению аппаратуры и расширению частотного диапазона, занимаемого системой, что не всегда возможно. Подход, рассмотренный в настоящей работе, лишен всех упомянутых недостатков. Это достигается использованием слепого метода разделения источников, в рамках которого в поле искусственных нейронных сетей формируется область возбуждения с обработкой сигнала.

Работа построена следующим образом. Общая структура предложенной системы представлена в разделе 2. В разделе 3 описаны альтернативные варианты устройства разделения сигналов, включая классические нейронные сети, а также новые, предложенные нами решения. В разделе 4 представлены экспериментальные результаты всех вариантов системы. Заключение дано в разделе 5, в приложениях рассмотрены некоторые специфические вопросы.

2. Предлагаемая система

Система, предлагаемая в данной работе (рис.2), предназначена для одновременной обработки сигналов двух меток и построена на основе описанной выше обычной системы.

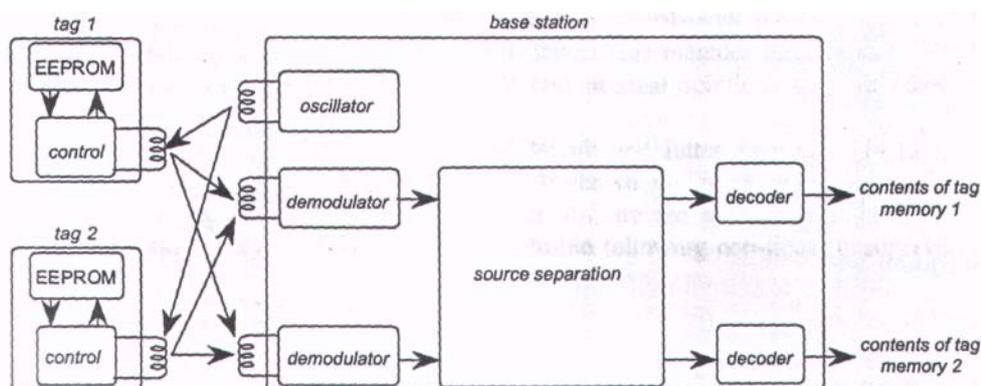


Рис.2. Система радиочастотной идентификации с несколькими метками.
tag – метка, base station – считыватель, oscillator – генератор, control – управление, demodulator – демодулятор, source separation - разделение сигналов, decoder – декодер, content of tag memory – содержимое памяти метки.

Она состоит из считывателя с двумя приемными антеннами и двумя демодуляторами, которые обрабатывают два смешанных сигнала. Эти сигналы обрабатываются устройством слепого разделения источников, которое выделяет два компонента, соответствующие двум меткам. После декодирования отдельных сигналов, образуется пара независимых идентификационных данных, которые хранятся в памяти каждой из двух меток.

В Приложении А описана теоретическая модель, положенная в основу обработки смеси сигналов, а результаты эксперимента, подтверждающие такую модель приведены в разделе 4.2. Показано, что использованные в системе устройства модуляции и демодуляции таковы, что демодулированная смесь сигналов может быть представлена в наиболее простом виде, т.е. является мгновенной линейной смесью (как определено в разделе 3) компонентов, соответствующих двум меткам. За последние 15 лет было предложено множество методов для разделения такой смеси сигналов. Обзор этих результатов можно найти в [1]. В данной работе мы используем подход, связанный с применением искусственных нейронных сетей, и мы исследуем их характеристики применительно к предложенной системе. Мы также рассмотрим модифицированную версию такого подхода, которая была предложена нами в последнее время, и мы сопоставим полученные в

данном случае результаты с результатами классического подхода. Все эти методы и подходы рассмотрены в разделе 3. В нашем исследовании они были выбраны по следующим причинам. Во-первых, свойства этих методов хорошо определены в рамках конфигурации двух источников и они применимы к типу рассматриваемых источников, что будет показано в последующих разделах данной работы. Кроме того, эти методы основаны на адаптивных алгоритмах, которые легко позволяют распространить полученные результаты на случай движущихся меток. И, наконец, они используют очень простые вычисления, что особенно привлекательно для получения недорогого конечного результата в реальном масштабе времени, являющегося целью настоящего исследования.

Необходимо отметить, что рассматриваемая система должна удовлетворять следующим требованиям, определенным в разделе 1:

1) она должна обладать высоким быстродействием, обеспечивая одновременную связь двух меток со считывателем; 2) все метки имеют одинаковое простое построение, как в обычной системе с одной меткой, и дополнительное усложнение допускается только в считывателе, т.е. с целью снижения достижения стоимости в единственном месте системы; 3) система использует единственную несущую частоту.

3. Проблема слепого разделения источников и ее решение

В простейшей постановке проблемы слепого разделения источников рассматриваются два неизвестных сигнала $E_1(t)$ и $E_2(t)$, которые являются линейной смесью двух неизвестных статистически независимых сигналов $X_1(t)$ и $X_2(t)$:

$$E_1(t) = a_{11}X_1(t) + a_{12}X_2(t), \quad (1)$$

$$E_2(t) = a_{21}X_1(t) + a_{22}X_2(t), \quad (2)$$

где a_{ij} – неизвестные действительные коэффициенты. Слепое разделение источников заключается в нахождении оценки сигналов $X_j(t)$ на основе анализа смеси сигналов $E_i(t)$ при произвольной перестановке и произвольных масштабных коэффициентах. Как отмечалось выше, основной проблемой, стоящей перед рассматриваемой системой, является ситуация, когда на основе анализа смеси сигналов на выходе демодуляторов, необходимо при помощи декодирования восстановить исходные данные, содержащиеся в памяти меток. В этом разделе описываются все решения этой основной проблемы, рассматриваемые в рамках данной работы. Они включают три однотипных известных нейронных решения и два оригинальных решения, полученных нами недавно.

Первый классический подход, предложенный в [14,15] и базирующийся на рекуррентной нейронной сети, показан на рис.3 для

случая описанной выше «простейшей конфигурации». Коэффициенты c_{12} и c_{21} являются адаптивными весами этой нейронной сети¹. Они определяются согласно следующему нелинейному правилу самообучения:

$$c_{ij}(n+1) = c_{ij}(n) - af[s_i(n)]g[s_j(n)], \quad (3)$$

где a – коэффициент позитивного адаптационного усиления, $s_i(t)$ и $s_j(t)$ – значения нейронных выходов $S_i(t)$ и $S_j(t)$ нейронной сети, а f и g – функции активации.

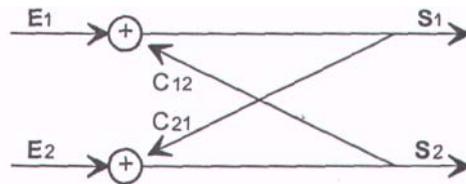


Рис.3. Рекуррентная нейронная цепь [14,15].

Функции f и g в основном зависят от вида сигналов, которые разделяет нейронная сеть. Их выбор исключительно важен для практического применения и может быть определен следующим образом. Когда используются произвольные нечетные функции f и g , нейронная сеть способна разделять некоторые типы симметричных сигналов [6]. Как показано в [6], это ограничение может быть исключено использованием $f = (\cdot)$ или $g = (\cdot)$ (но не обе одновременно, так как это приведет к использованию статистик сигнала только второго порядка и не гарантирует, что этот алгоритм обеспечит разделение [15]). Благодаря простоте и виду сигналов, к которым они применимы, особенно привлекательны два типа функций:

$$f = (\cdot)^3 \text{ и } g = (\cdot), \quad (4)$$

$$f = (\cdot) \text{ и } g = (\cdot)^3. \quad (5)$$

Выбор между этими двумя типами функций зависит от конкретного вида сигналов (для уверенности, что весовые коэффициенты нейронной сети сойдутся к значениям, которые обеспечивают разделение сигналов на выходе нейронной цепи): функции (4) используются для обобщенных sub-Гауссовых процессов [5,11,17,18,21], т.е. для сигналов, у которых $R < 9$, где R – отношение, определенное следующим образом:

$$R = E\{x_1^4\}E\{x_2^4\} / (E\{x_1^2\})^2 (E\{x_2^2\})^2, \quad (6)$$

¹ По сравнению с оригинальными работами Herault and Jutten, для удобства последующего изложения знаки весов c_{12} и c_{21} на рис.3 изменены. Соответственно модифицированы веса в формуле (3).

где $x_j(t)$ – центрированное значение сигнала $X_j(t)$, а $E\{\}$ – математическое ожидание. Можно показать¹, что функции (5) применимы к обобщенным super-Гауссовым процессам, т.е. к процессам с $R > 9$.

В Приложении В теоретически показано, что для рассматриваемой системы процессы являются обобщенными sub-Гауссовыми. Поэтому вид цепи, которая предназначена для разделения таких сигналов, должен соответствовать функциям (4). Экспериментально это подтверждено в разделе 4.

В работах [16–18] предложен прямой (не рекуррентный) вариант нейронной сети (см. рис.4), функционирующий по тому же правилу (3), как и для сети, представленной на рис.3. Авторы работ [16–18] также рассмотрели вопрос сходимости свойств этой сети и показали, что в этом случае использование функций (4) обеспечивает разделение sub-Гауссовых процессов.

В работе [2] также предложена нейронная сеть, которую можно рассматривать как развитие сетей, рассмотренных ранее. Эта сеть содержит дополнительные самонастраивающиеся веса, которые настраиваются так, чтобы нормировать "масштаб" выходов сети. В этой работе описаны прямая и рекуррентная версии такого типа нейронных сетей, а также для улучшения характеристик предложено каскадировать их в многослойную нейронную сеть. В настоящей работе мы рассмотрим только прямую версию, так как она обладает такими же свойствами, как сеть, рассмотренная в [16–18], по сравнению с соответствующей рекуррентной структурой. Более того, мы остановимся на однослойном варианте такой сети. В нашей работе [7] показано, что для таких сетей функции (4) также обеспечивают разделение sub-Гауссовых сигналов.

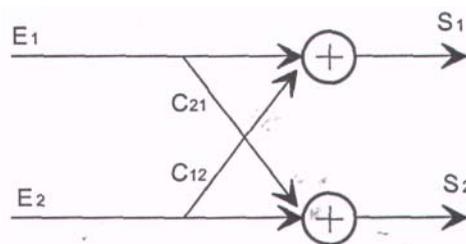


Рис.4. Нейронная сеть с последовательными связями.

Мы также рассмотрим другой вариант самонормирующихся нейронных сетей разделения сигналов, недавно предложенный нами. Принцип действия и особенности этих цепей подробно рассмотрены в [9] и представлены далее. Эти сети базируются на основе тех же самых классических структур, которые были рассмотрены выше: они могут иметь один или несколько слоев и каждый слой может быть в рекуррентной или

¹ С использованием функций (5) при адаптации метода [21].

прямой форме. Однослойные версии таких сетей с двумя источниками сигнала представлены на рис.3 и рис.4 соответственно, а их расширение на большее число источников сигнала осуществляется аналогично сетям, предложенным в работах [14–18]. Предложенные сети отличаются от ранее рассмотренных алгоритмом вычисления весов:

$$c_{ij}(n+1) = c_{ij}(n) - a (f[s_i(n)]/(E[f^2(s_i)])^{1/2})(g[s_j(n)]/(E[g^2(s_j)])^{1/2}). \quad (7)$$

Нормированные величины $(E[f^2(s_i)])^{1/2}$ и $(E[g^2(s_j)])^{1/2}$, введенные нами, по сравнению с классическим правилом (3) получаются путем низкочастотной фильтрации первого порядка. Когда функции f и g удовлетворяют (4), правило адаптации (7) приобретает следующий вид:

$$c_{ij}(n+1) = c_{ij}(n) - a(s_i^3(n)/\sqrt{E[s_i^6]})(s_j(n)/\sqrt{E[s_j^2]}). \quad (8)$$

Когда в каждый из сигналов смеси $s_i(t)$ и $s_j(t)$ вводится масштабный коэффициент λ , правая часть (8) благодаря введенной нами нормировке не изменяется. Весовые траектории, достигаемые согласно этому правилу, и, следовательно, скорость и точность сходимости предложенных сетей полностью нечувствительны к общему уровню сигналов, что весьма привлекательно из-за неопределенности уровня на практике. Скорость и точность сходимости могут изменяться выбором адаптационного усиления a : низкое усиление обеспечивает хорошую точность сходимости за счет низкой скорости сходимости, а высокое усиление приводит к обратному эффекту, так что выбором величины адаптационного усиления можно обеспечить желаемое соотношение скорость/точность сходимости. Напротив, классические сети [14–18] не способны обеспечивать контролируемое соотношение скорость/точность сходимости. Это основной мотив для рассмотрения предложенного нами самонормирующегося правила. Такое правило имеет несколько других привлекательных свойств, большинство из которых сохраняется при других видах функций f и g . Полное описание этих свойств выходит за рамки настоящей работы, но некоторые из них могут быть найдены в [9]. Кроме того, остаются верными те же рассуждения по поводу выбора функций f и g , которые приводились выше, т.е. возможен выбор между (4) и (5) в зависимости от вида обрабатываемого процесса (более подробно см. в [9]).

4. Экспериментальные результаты

4.1. Экспериментальная установка

Экспериментальная установка, используемая для проверки эффективности предложенного подхода, представлена на рис.5. Антенны и

метки представляют собой горизонтальные диски (диаметром 52 мм для антенн считывателя и 28 мм для меток). Метки размещаются на горизонтальной пластмассовой пластине; антенны также размещаются в горизонтальной плоскости на расстояниях h_1 и h_2 ниже плоскости меток. Как указывалось выше, при проведении данного исследования использовались элементы стандартной коммерческой системы радиочастотной идентификации: антенны, демодуляторы и метки. Дальность действия такой системы ограничена, т.е. метки для уверенного функционирования должны располагаться на расстоянии не более 60 мм от антенн считывателя. В рассматриваемой установке это требование обеспечением расстояний $h_1 = 35$ мм и $h_2 = 25$ мм. Расстояние D между метками варьировалось в процессе эксперимента. Когда метки располагаются близко одна от другой (т.е. D близко к диаметру метки), стандартная система не может идентифицировать метки, поэтому необходимо устройство разделения сигналов. Такая ситуация и рассматривается в данной работе.

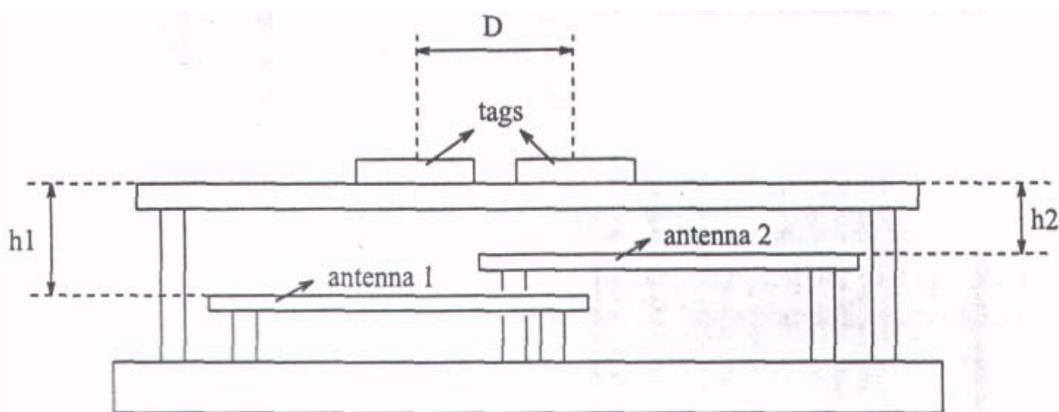


Рис.5. Вертикальное сечение экспериментальной установки.

tags – метки, antenna – антенна.

Следует заметить, что размеры, использованные в установке, типичны для различных применений системы, как, например, для автомобильных иммобилайзеров. В этом случае метка встроена в обычный механический ключ и ее содержимое проверяется, когда владелец автомобиля запускает двигатель. Ситуация с несколькими метками возникает в случае, когда один или несколько электронных ключей, предназначенных для других применений (для другого автомобиля, например), находятся близко к ключу рассматриваемого автомобиля, так, что сигналы других ключей искажают сигнал нужного ключа. В этом случае обычный считыватель с одной антенной не способен идентифицировать сигнал ключа, принадлежащего данному автомобилю, и требуется система, подобная той, которая обсуждается в настоящей работе. Некоторые применения, упомянутые в

разделе 1, требуют расстояний, значительно больших, чем те, которые рассматриваются в настоящей работе. В этом случае наша установка может рассматриваться как маломасштабная модель необходимой системы. Такая система может быть преобразована из нашей установки заменой приемопередающего устройства, которое обеспечит большую дальность.

4.2. Экспериментальные свойства источников и смесей сигналов

Первые эксперименты, проведенные на описанной выше установке, обеспечили проверку того, что сигналы и смеси соответствуют теоретической модели, изложенной в предыдущем разделе работы.

Первый ряд экспериментальных измерений показал, что каждый из сигналов $x_j(t)$ и отношение $E\{x_j^4\}/(E\{x_j^2\})^2$ близки к 1,3. Поэтому параметр R , определенный в (6), для каждой пары сигналов близок к 1,7. Это подтверждает, что реальные сигналы являются обобщенными sub-Гауссовыми. Поэтому все эксперименты по разделению сигналов, приведенные в следующем подразделе, выполнялись с использованием нейронных сетей и функций, определенных согласно (4).

Затем экспериментально проверялись свойства смесей сигналов. Как и в начале, обеспечивалась проверка того, что смеси одновременны. Для этого мы использовали два метода. Во-первых, была проведена качественная верификация размещением единственной метки в поле считывателя с двумя антеннами и визуальным анализом выходных сигналов обоих демодуляторов на экране осциллографа. Эти два сигнала были пропорциональными с нулевой временной задержкой, что соответствует одновременной смеси (см. (1) – (2): два сигнала, соответствующие одной метке, должны быть равными $a_{11}X_1(t)$ и $a_{21}X_1(t)$). Затем этот результат был подтвержден количественным методом: с частотой 32 кГц были сделаны выборки двух выходных сигналов демодуляторов¹. Затем был проанализирован кросскорреляционный коэффициент этих двух сигналов:

$$\rho(\tau) = \{s_1(t)s_2(t+\tau)\}/(\sqrt{E\{s_1^2(t)\}})(\sqrt{E\{s_2^2(t)\}}). \quad (9)$$

¹ Эта частота была выбрана только потому, что в рассматриваемой установке она была единственно возможной. Вероятно, для обработки сигналов может быть использована более низкая частота.

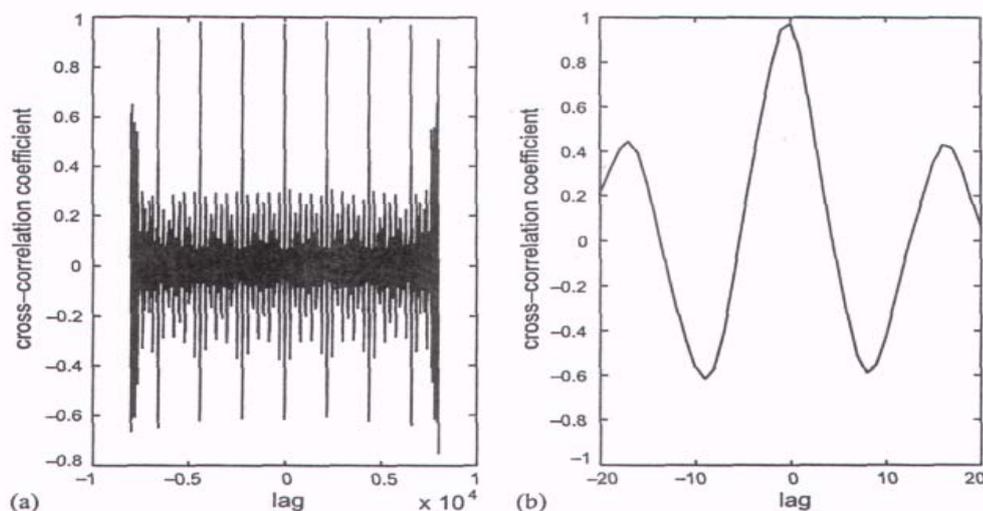


Рис.6. Коэффициент кросскорреляции: (а) - полная последовательность
(б) - в окрестности нуля.

cross-correlation coefficient – коэффициент кросскорреляции.

На рис.6а представлен кросскорреляционный коэффициент, полученный по полному объему выборки, который равен 8000. Периодические свойства коэффициента обусловлены периодичностью сигналов, так как они состоят из одинаковых фрагментов, содержащих около 2000 отсчетов (см. Приложение В). Увеличенный масштаб кросскорреляционного коэффициента вблизи точки 0 показан на рис.6б. Здесь видно, что коэффициент кросскорреляции $\rho(\tau)$ достигает максимального значения точно при $\tau = 0$, а его максимальное значение близко к 1. Поэтому два рассматриваемых сигнала пропорциональны друг другу с нулевой задержкой, что подтверждает одновременные свойства смеси сигналов.

Второй шаг нашего исследования состоял в экспериментальной проверке линейности смеси сигналов. Основными источниками нелинейности рассматриваемой системы, вероятно, являются демодуляторы, тогда как совокупность сигналов, принятых антеннами, предполагается линейной. Поэтому анализировалась линейность только демодуляторов. Для этого на вход демодулятора подавался сигнал на несущей частоте, модулированный синусоидой. Мы варьировали величину модулирующей синусоиды и анализировали выходной сигнал демодулятора. Результаты представлены на рис.7. Видно, что демодулятор остается линейным при амплитуде модулирующего сигнала, не превышающей 1,5 В. Отсюда было определено минимальное расстояние между метками и антеннами, гарантирующее линейное функционирование системы, которое по величине близко к 1 см.

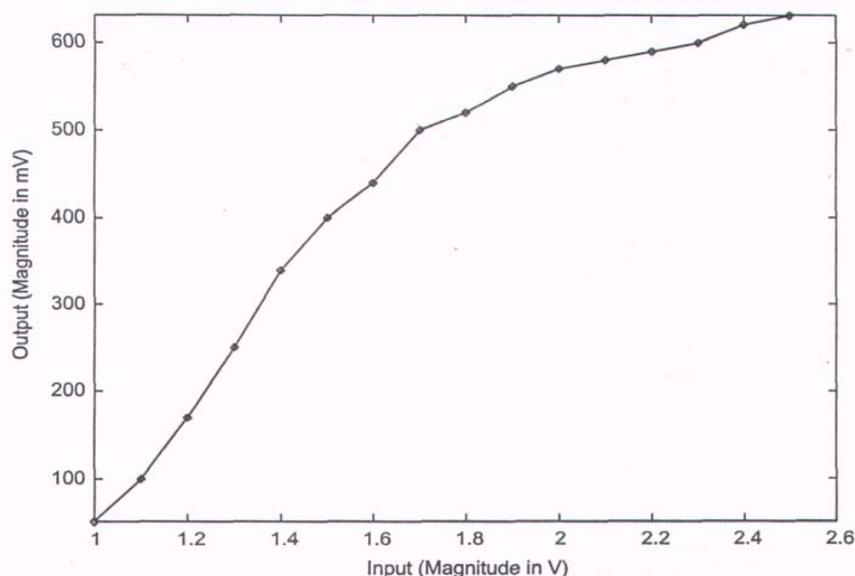


Рис.7. Зависимость сигнала на выходе демодулятора от амплитуды модулирующего сигнала.

Output (Magnitude in mV) – Выходной сигнал, мВ, Input (Magnitude in V) – Входной сигнал, В.

Вот почему, как описано ранее, метки в рассматриваемой системе были расположены на плоскости, отстоящей на несколько сантиметров выше антенн.

Все исследования, результаты которых приведены в настоящем подразделе, показывают, что источники и смеси сигналов в рассматриваемой установке обладают свойствами, требуемыми для применения нейронных сетей. С учетом этого мы продолжили проведение экспериментов, сфокусировав внимание на характеристиках системы, которые касаются разделения входных сигналов.

4.3. Разделение искусственных смесей сигналов

Первый этап в эксперименте по разделению сигналов был выполнен с искусственной смесью реальных сигналов, поступающей на вход каждой из пяти типов нейронных цепей, рассмотренных в разделе 3:

- нейронная сеть [14, 15];
- нейронная сеть [16, 17, 18];
- однослойная нейронная сеть с последовательными связями[2];
- две однослойных нейронных сети, основанные на рекуррентной и последовательной структуре, ниже обозначенные NWU_r и NWU_d соответственно (где "NWU" означает нормированный весовой алгоритм самообучения, использованный в этих цепях).

Цель этих экспериментов была двойкой. С одной стороны, они помогли проверить, что все эти цепи действительно способны разделять сигналы в реальной системе в предположении, что смесь является линейной и одновременной. С другой стороны, они позволили сравнить характеристики всех рассматриваемых цепей в различных условиях и выбрать лучшую из них.

Более точно, эксперименты были выполнены при следующих условиях. Для того чтобы образовать линейную одновременную смесь реальных сигналов, сначала одна первая метка была размещена в радиочастотном поле установки. С целью получения первого сигнала $X_1(t)$ была сделана выборка выходного сигнала одного из демодуляторов. Затем эта метка была удалена и в поле считывания была размещена вторая метка. Для получения второго сигнала $X_2(t)$ операция, аналогичная предыдущей, была выполнена со второй меткой. Затем, в соответствии с (1) – (2), были вычислены две искусственные смеси $E_1(t)$ и $E_2(t)$. Эти смеси были подвергнуты обработке на вычислительной машине с плавающей запятой по соответствующим нейросетевым алгоритмам. Исследовались два случая коэффициентов смеси a_{ij} . В обоих случаях a_{11} и a_{22} принимались равными 1. Сложность состояла при выборе значений a_{12} и a_{21} , которые были определены следующим образом:

- первые эксперименты были выполнены при $a_{12} = 0,4$ и $a_{21} = 0,3$. Эти величины соответствуют среднему значению отношения смеси и соответствуют действительным значениям, характерным для смеси сигналов в экспериментальной установке (их можно было вычислить из результатов, изложенных в разделе 4.4).
- другие эксперименты были выполнены при $a_{12} = a_{21} = 0,98$. Эти величины соответствуют очень высокому значению отношения смеси, которое характерно для систем радиочастотной идентификации большой дальности, при этом две метки очень близки одна к другой по сравнению с расстоянием до антенн. Можно предположить, что такие сигналы очень трудно делимы, так как, исходя из подстановки этих значений a_{ij} в (1) – (2), смеси $E_1(t)$ и $E_2(t)$ очень похожи.

Характеристики, достигнутые в каждом эксперименте, сравнивались по двум параметрам, рассмотренным в разделе 3, т.е. по скорости и точности сходимости исследуемой нейронной сети. Скорость сходимости оценивалась по числу выборок, которое требуется всем весовым коэффициентам нейронной цепи для сходимости к их устойчивым значениям, и называется временем сходимости¹, обозначенным в дальнейшем T_c . Точность сходимости оценивалась улучшением отношения сигнал/шум ($SNRI$), которое достигается рассматриваемой нейронной сетью. Этот параметр, принимающий большие значения, когда сеть

¹ Времена сходимости оценивались из графиков, показывающих эволюцию весов цепи во времени.

восстанавливает хорошо разделенные сигналы, определяется следующим образом: $SNRI = (SNRI_1 + SNRI_2)/2$. Каждый из параметров $SNRI_i$ в этом выражении обозначает улучшение отношения сигнал/шум, обеспечиваемое на i -ом выходе сети по отношению к i -му сигналу.

Для нейронных сетей, рассмотренных в [14, 15], и NWUr сетей, когда разделение сигналов без перестановки достигается точно, $S_i(t)$ становится равным $a_{ii}X_i(t)$ [8,15]. Поэтому для этих нейронных сетей $SNRI_i$ равно:

$$SNRI_i = 10\log_{10}[(E\{(E_i(t) - a_{ii}X_i(t))^2\})/(E\{(S_i(t) - a_{ii}X_i(t))^2\})]. \quad (10)$$

Аналогично для нейронных сетей, рассмотренных в [16, 17. 18], и NWUd сетей:

$$SNRI_i = 10\log_{10}[(a_{11}a_{22} - a_{12}a_{21}/a_{11}a_{22})^2(E\{(E_i(t) - a_{ii}X_i(t))^2\})/(E\{(S_i(t) - ((a_{11}a_{22} - a_{12}a_{21}) X_i(t))/a_{ij})^2\})], \quad (11)$$

тогда как для нейронной сети, рассмотренной в работе [2]:

$$SNRI_i = 10\log_{10}[(1/a_{ij}^2 E\{x_i^4(t)\}^{1/2})(E\{(E_i(t) - a_{ii}X_i(t))^2\})/E\{(S_i(t) - (X_i(t)/E\{x_i^4(t)\}^{1/4})^2\})]. \quad (12)$$

Как указано в разделе 3, обобщенной характеристикой данных нейронных сетей является соотношение между T_c и $SNRI$, достигаемое этими сетями. Это соотношение было определено выполнением эксперимента для различных значений коэффициента адаптации a , регистрацией T_c и $SNRI$, которые получены при этих условиях, и графическим отображением результирующих изменений $SNRI$ во времени T_c . Полученные результаты показаны на рис.8. и рис.9. для двух упомянутых выше значений отношения смеси соответственно.

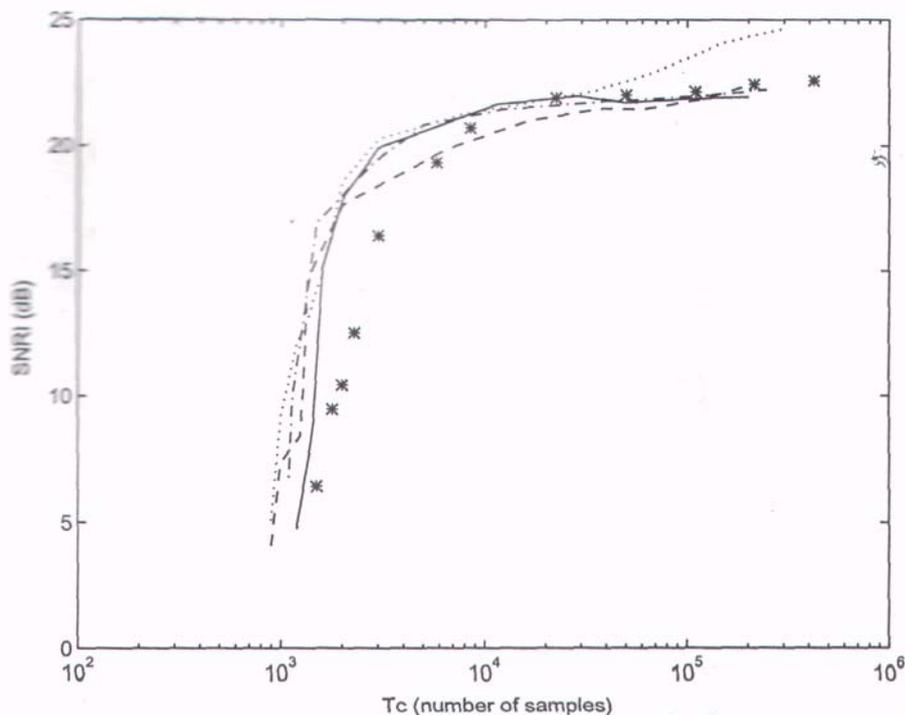


Рис.8. Зависимость $SNRI$ от времени сходимости T_c при $a_{12} = 0,4$ и $a_{21} = 0,3$. Нейронные сети: [14, 15] -.-.-; [16, 17, 18]; [2] ***; NWUr — и NWUd ----. number of samples – число отсчетов.

Наиболее интересная часть графиков, соответствующая диапазону изменения T_c , который встречается на практике, может быть определена следующим образом. Как отмечалось ранее, данные, которые поступают от метки на считыватель в стандартной системе радиочастотной идентификации с одной меткой, состоят из серии идентичных фрагментов, содержащих около 2000 отсчетов. Кроме того, когда метка оказывается в радиочастотном поле и начинает функционировать, считыватель находится в режиме ожидания до момента приема синхронизирующей последовательности (что соответствует началу фрагмента), после чего начинает декодировать принятый сигнал. Другими словами, считыватель имеет внутреннюю задержку длительностью приблизительно в один фрагмент. Поэтому в системе с несколькими метками желательно, чтобы для необходимости обеспечения разделения сигналов в первом принятом полном информационном фрагменте время сходимости алгоритма не превышало времени внутреннего ожидания. Тогда введение устройства разделения в обычную идентификационную систему с одной меткой не снизит ее быстродействия. Следовательно, типичное значение приблизительно равно длительности одного фрагмента, т.е. около 2000 отсчетов. Более того, различные применения могут допускать и большее время задержки (обычно несколько фрагментов), так как длительность одного фрагмента около 70 мс, а это является достаточно небольшим временем по сравнению с требуемым временем реакции системы во многих

идентификационных приложениях. Поэтому оценка характеристик рассматриваемых цепей производилась не только при значении $T_c \approx 2000$ отсчетов, но также в диапазоне $T_c = 2000 - 10000$ выборок (т.е. до 5 фрагментов).

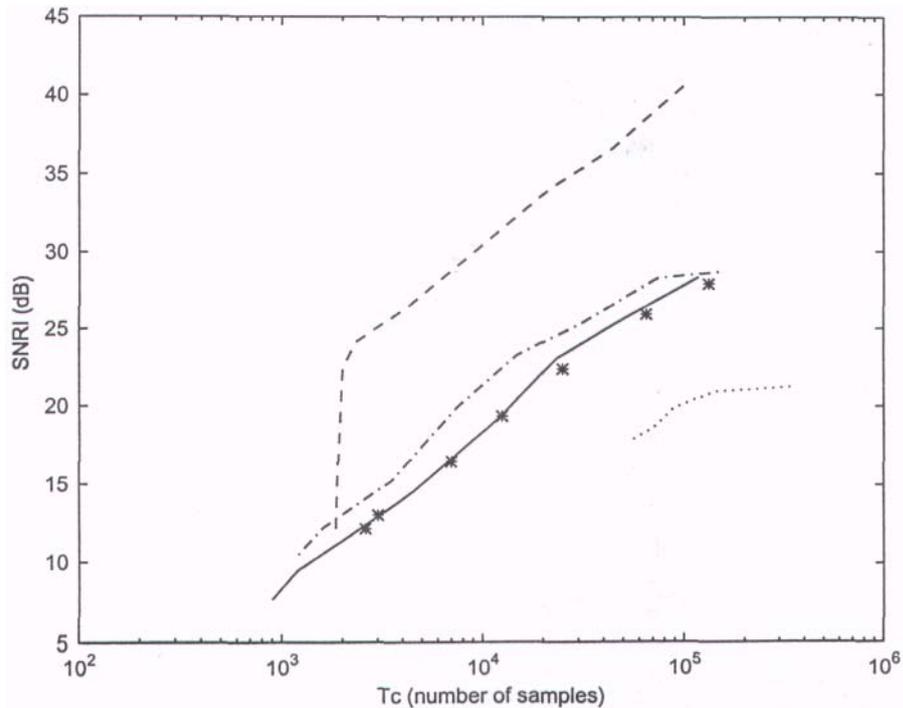


Рис.9. Зависимость $SNRI$ от времени сходимости T_c при $a_{12} = a_{21} = 0,98$. Нейронные сети: [14, 15] -.-.-; [16, 17, 18]; [2] ***; NWUr — и NWUd - - - (для сетей [16, 17, 18] и [2] значения T_c меньше, чем приведенные на этом рисунке, не могут быть получены, так как T_c и $SNRI$ становятся очень чувствительными к увеличению коэффициента адаптации a и эти сети возможно расходятся).

Рис.8 и 9 показывают, что нейронная сеть [16, 17, 18] в рассматриваемом случае скорее всего не может использоваться, из-за того, что она позволяет получить нужного T_c для больших значений отношения смеси.

Сеть [2] не привлекательна по двум причинам: (i) она не достигает $T_c \cong 2000$ отсчетов (или ее $SNRI$ слишком мал) и (ii) для любых T_c интересующего нас диапазона ее $SNRI$ ниже, чем для остальных трех типов нейронных сетей, т.е. для сетей [14, 15], NWUr и NWUd сетей. Среди последних трех типов нейронных сетей предпочтение зависит от основного параметра, который важен для конкретного применения системы. Все три типа нейронных сетей достигают $T_c \cong 2000$ отсчетов, но это значение почти предельно для NWUd сети. Поэтому, если наиболее важным является минимизация T_c , нейронные

сети [14, 15] и NWUг сети предпочтительны. И наоборот, если акцент делается на использование с большим значением отношения смеси, а параметр T_c не столь критичен (в пределах рассматриваемого диапазона), предпочтительной является NWUd нейронная сеть.

До этого момента мы рассматривали только характеристики T_c и $SNRI$ нейронных сетей. Однако, как пояснялось в разделе 3, следует учитывать и другие свойства, т.е. их способность работать в самонормирующемся («автоматическом») режиме, или, что то же самое, быть независимыми от уровня входного сигнала. По этой причине, кроме [16, 17, 18] и [2] нейронных сетей, приходится отказаться и от [14, 15] нейронных сетей. Другими словами в рассматриваемых случаях предпочтительными оказываются NWUг и NWUd сети. Окончательный выбор, как отмечалось выше, можно делать в зависимости от акцента на T_c или на значение отношения смеси. Поэтому в дальнейшем рассматриваются только эти два типа нейронных сетей. Кроме того, их коэффициент усиления адаптации a устанавливается таким, чтобы обеспечить $T_c \cong 2000$ отсчетов в экспериментах с коэффициентами смеси, упомянутыми ранее, т.е. $a = 10^{-3}$ для обоих типов нейронных сетей.

4.4. Разделение реальных смесей

Вторая группа экспериментов по разделению сигналов была выполнена на реальной системе. Две метки одновременно были помещены в радиочастотное поле и были измерены две смеси $E_1(t)$ и $E_2(t)$ на выходе обоих демодуляторов. Эти две реальные смеси использовались в качестве входных сигналов вычислительных NWUг и NWUd нейронных сетей. На рис.10 показана эволюция весов NWUг сети при величине обучающего усиления $a = 10^{-3}$. Это усиление обеспечивает $T_c \cong 2000$ отсчетов, что полностью совпадает с результатами, полученными для искусственных смесей в разделе 4.3. Эксперименты, выполненные с NWUd сетью, привели к таким же результатам.

Рис.10 также показывает, что весовые коэффициенты нейронных сетей сходятся к различным значениям. Это объясняется тем, что величины весов, для которых производилось разделение сигналов, зависят от величин коэффициентов смеси [8], которые различны вследствие физической асимметрии установки (см. рис.5).

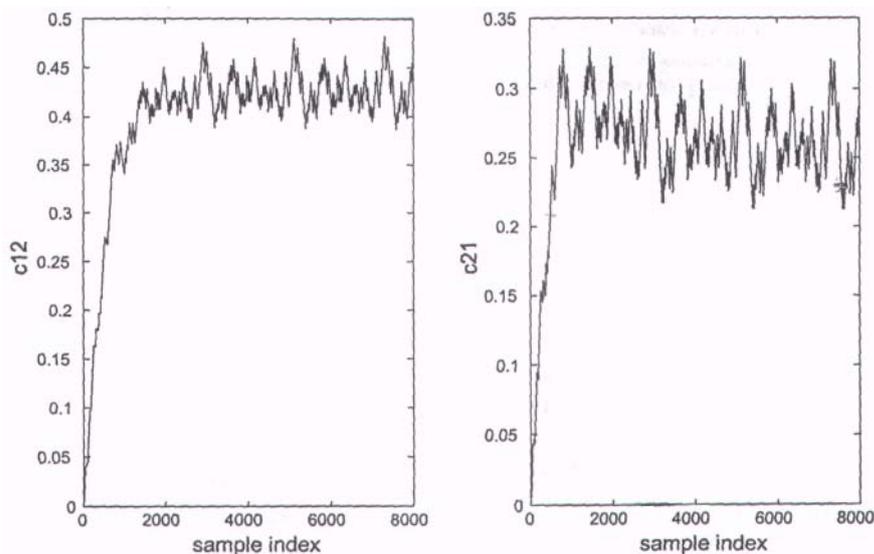


Рис.10. Эволюция весов нейронной сети NWU Γ для реальной смеси.
sample index – номер отсчета.

Поскольку коэффициенты смеси неизвестны, теоретические величины весов, соответствующие разделению сигналов, и экспериментальные $SNRI$ (10) – (11) не могут быть вычислены. Поэтому для проверки возможности восстановления исходных сигналов был использован другой подход. Использованный здесь альтернативный метод заключается в подключении выходов нейронных сетей к входам декодеров системы. Как описывалось выше, эти декодеры находятся в режиме ожидания первого синхронизирующего пакета на выходе сети и затем обеспечивают восстановление данных, которые записаны в метках. Сравнение полученных данных с оригинальными данными меток¹ показали, что они полностью совпадают.

Таким образом,

- 1) нейронные сети NWU Γ и NWU Δ не снижают быстродействие системы, так как время их сходимости примерно равно длительности одного фрагмента данных, в течение которого декодеры в любом случае должны ожидать поступления синхронизирующего пакета, и
- 2) после сходимости они обеспечивают полное восстановление сигналов в том смысле, что они восстанавливают идентификационные номера, хранящиеся в метках, без каких-либо ошибок.

4.5. Практическая реализация

Кроме успешно проведенных экспериментов, мы, на основе описанных выше подходов и с использованием DSP устройства, которое работало только с фиксированной запятой, разработали реальную установку.

¹ Конечно, в нашем случае данные, хранящиеся в метках, были известными, однако на практике при функционировании идентификационных систем, они неизвестны.

В результате большого числа экспериментов мы показали, что скорость сходимости такой установки почти такая же, как у системы с плавающей запятой, и что данные меток декодируются сразу после обеспечения сходимости. Эти опыты также показали долговременную стабильность рассматриваемых нейронных сетей. Более того, они позволили нам проверить возможности таких сетей по отслеживанию изменяющихся смесей.

5. Обсуждение и заключение

Исследования, представленные в настоящей работе, демонстрируют, что нейронные сети, предназначенные для слепого разделения сигналов, позволяют обеспечить функционирование системы радиочастотной идентификации при ее незначительной доработке. Среди всех рассмотренных методов и устройств наиболее привлекательными оказались те, которые использовали предложенный нами подход, благодаря простоте, хорошим характеристикам и самонормирующимся свойствам.

Заслуживают комментариев возможные направления развития данного исследования. Предложенные подходы наиболее привлекательны в рассмотренной конфигурации системы, так как для данного случая применения системы требуются малые вычислительные ресурсы, тогда как большинство других применений потребует более сложной аппаратуры обработки сигналов. Некоторые из применений, известных из литературы, отмечены ниже:

- в будущих применениях идентификационных систем, которым необходимо высокое быстродействие и, следовательно, разделение сигналов с малым временем сходимости;
- и в применениях, содержащих более двух меток. Для случаев, когда необходимо разделять более чем два сигнала, развитие рассматриваемых в настоящей работе подходов может быть легко найдено. При этом, однако, на данной стадии работы свойства сходимости недостаточно изучены. Альтернативой может быть использование некоторых подходов с гарантированно известным временем сходимости. Это, в частности, включает подход, предложенный в [3] и развитый в нескольких вариантах в работе [13].

В любом случае необходимо помнить, что различные идентификационные системы имеют ограниченную зону считывания. В этих системах лишь метки, расположенные вблизи считывателя, обеспечиваются достаточным сигналом. Это в принципе ограничивает количество меток до небольшого количества и даже, возможно, до двух штук. Поэтому применения с большой дальностью идентификации также могут использовать конфигурацию системы и решения, рассмотренные в данной работе.

Другое возможное развитие настоящего исследования касается использования разделения сигналов для уменьшения переотражений от окружающих предметов, что позволяет

- (i) увеличить расстояние между считывателем и метками или
- (ii) уменьшить энергопотребление.

Возможно, также, что в рассмотренном методе были лишь частично использованы априорные сведения о сигналах. Это позволило нам разработать метод, который может быть развит на другие идентификационные системы. Также может быть разработан метод точной настройки с использованием устройства разделения сигналов, который применен в специфической системе, рассмотренной в данной работе, и который способен обеспечить более эффективную обработку сигналов.

В заключение авторы благодарят А. Laurence за поддержку на начальной стадии работы, а также анонимных рецензентов за их отзывы.

Приложение А. Теоретические свойства смесей

В этом приложении показано, что в идеальной модели системы, рассмотренной в разделе 2, смеси сигналов, подвергающиеся обработке устройством разделения сигналов, являются линейной одновременной смесью компонентов, соответствующих двум меткам. Как показано ниже, эти свойства смесей вытекают из схемы модуляции/демодуляции, использованной в данной работе, т.е. низкочастотной амплитудной модуляции.

Для доказательства этого рассмотрим процесс распространения и обработки сигналов. Сначала несущий сигнал $V \sin \omega t$ излучается считывателем. Он принимается каждой меткой j с коэффициентом α_j (связанным с затуханием сигнала в канале распространения) и задержкой τ_j . Тогда сигнал, принятый меткой j в момент t , будет равен $\alpha_j V \sin [\omega(t - \tau_j)]$. Предположим, что метка осуществляет идеальную амплитудную модуляцию несущей кодовым сигналом $x_j(t)$, тогда сигнал, излученный меткой j в момент t , будет равен $x_j(t) \alpha_j V \sin [\omega(t - \tau_j)]$.

Теперь рассмотрим одну из приемных антенн считывателя (рассуждения верны для любой из антенн). Она принимает каждый излученный меткой сигнал с коэффициентом α'_j и задержкой τ'_j . Сигнал, принятый этой антенной в момент t , будет равен $a_j x_j(t - \tau'_j) V \sin [\omega(t - \tau_j - \tau'_j)]$, где $a_j = \alpha_j \alpha'_j$. Антенна идеально суммирует сигналы, принятые от двух меток. Соответствующий выходной сигнал антенны в момент t , будет равен:

$$a_1 x_1(t - \tau'_1) V \sin [\omega(t - \tau_1 - \tau'_1)] + a_2 x_2(t - \tau'_2) V \sin [\omega(t - \tau_1 - \tau'_2)]. \quad (\text{A.1})$$

Этот сигнал можно записать в следующем виде:

$$\begin{aligned} & \sin \omega t [a_1 V x_1(t - \tau'_1) \cos \varphi_1 + a_2 V x_2(t - \tau'_2) \cos \varphi_2] + \\ & + \cos \omega t [a_1 V x_1(t - \tau'_1) \sin \varphi_1 + a_2 V x_2(t - \tau'_2) \sin \varphi_2], \end{aligned} \quad (\text{A.2})$$

где $\varphi_j = -\omega(\tau_j + \tau'_j) = -2\pi f L_j / c$, f – несущая частота, L_j – общее (удвоенное) расстояние между антенной и меткой, $\tau'_j = l_j / c$ – расстояние между антенной и меткой.

В рассматриваемой системе параметры имеют следующие числовые значения: $f = 125$ кГц и $L_j < 10$ см, так что $\varphi_j < 2,6 \times 10^{-4} \ll 1$, $l_j < 5$ см, так что $\tau'_j < 1,7 \times 10^{-10}$ с $\ll T \cong 10^{-3}$ с, где T – период модулирующего сигнала.

Используя эти значения параметров, суммарный сигнал, принятый антенной, можно аппроксимировать следующим образом: $\sin \omega t [a_1 V x_1(t) + a_2 V x_2(t)]$. Результирующий идеальный амплитудно демодулированный сигнал будет равен $a_1 V x_1(t) + a_2 V x_2(t)$. Этот сигнал, который подается на вход устройства разделения сигналов, является линейной одновременной смесью двух модулирующих сигналов $x_1(t)$ и $x_2(t)$ (как показано в разделе 3). Следует заметить, что этот результат применим в случае, когда сигнал каждой метки распространяется до считывателя несколькими путями: каждый путь вносит свой индивидуальный вклад. Однако, когда можно пренебречь задержками распространения, все эти вклады суммируются в общий сигнал, соответствующий описанной модели.

Приложение В. Теоретические свойства сигналов

В этом Приложении мы представим используемую модель сигналов, которая должна быть восстановлена устройством разделения сигналов в системе. Каждый из сигналов состоит из одинаковых фрагментов. Каждый фрагмент содержит последовательность, синхронизирующую данные (т.е. содержимое памяти метки). Кроме того, эти данные кодируются с помощью стандартной бифазной кодирующей процедуры, которую мы рассмотрим ниже.

Сначала рассмотрим идеальное функционирование системы с точки зрения кодированных данных (т.е. без учета влияния синхронизирующей последовательности). В идеале каждый бит данных, равный 0, соответствует напряжению $+V$ в течение первой половины цикла и противоположному напряжению $-V$ в течение второй половины цикла. Бит, равный 1, кодируется изменяющимися значениями: напряжение $+V$ в течение полного цикла и напряжение $-V$ в течение следующего полного цикла. Такой идеальный сигнал представляет собой случайный стационарный сигнал, принимающий значения -1 и $+1$ с вероятностью $1/2$ независимо от значений кодируемых битов. В результате можно показать, что для пары разделяемых

сигналов $R = 1$, где R определяется (6). Такое значение R значительно меньше порогового уровня $R = 9$, т.е. сигнал является обобщенным sub-Гауссовым (см. раздел 3).

В реальной идентификационной системе сигналы существенно искажаются и не имеют бинарного вида. К тому же они содержат синхронизирующую последовательность, которая не симметрична. Соответствующее отношение R может отличаться от теоретического значения $R = 1$, но остается значительно меньшим порогового уровня $R = 9$. Таким образом, можно предположить, что реальные сигналы являются обобщенными sub-Гауссовыми. Это экспериментально подтверждено в разделе 4.

Список литературы

- [1] J.-F. Cardoso, Blind signal separation: statistical principles, Proc. IEEE 86 (1998) 2009–2025.
- [2] A. Cichocki, W. Kasprzak, S.I. Amari, Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals, Proceedings of the NOLTA '95, International Symposium on Nonlinear Theory and Its Applications, Las Vegas, 1995, pp. 61–65.
- [3] N. Delfosse, P. Loubaton, Adaptive blind separation of independent sources: a deflation approach. Signal Process. 45 (1995) 59–84.
- [4] P.R.M. Denne, C.D. Hook, Identification systems, UK patent no. 2 157 132 A, 16 October 1985.
- [5] Y. Deville, A unified stability analysis of the Herault-Jutten source separation neural network, Signal Process. 51 (1996) 229–233.
- [6] Y. Deville, Application of the Herault—Jutten source separation neural network to multi-tag radio-frequency identification systems, Proceedings of the Ecole des Techniques Avancees en Signal Image Parole, Grenoble, 1996, pp. 265–272.
- [7] Y. Deville, Analysis of the convergence properties of self-normalized source separation neural networks, IEEE Trans. Signal Process. 47 (1999) 1272–1287.
- [8] Y. Deville, N. Charkani, Analysis of the stability of time-domain source separation algorithms for convolutively mixed signals, Proceedings of the ICASSP 97, IEEE International Conference on Acoustics, Speech, and Signal Processing, IEEE Press, Munich, 1997, pp. 1835–1838.
- [9] Y. Deville, N. Charkani, Convergence of source separation neural networks operating with self-normalized weight updating terms, Proceedings of the ICA'99, International Workshop on Independent Component Analysis and Blind Signal Separation, Aussois, France, 1999, pp. 227–232.

- [10] S. Drews, W. Tobergte, V. Timm, K. Axer, Verfahren zum Steuern des Zugriffs auf einen Speicher sowie Anordnung zur Durchfuehrung des Verfahrens, German patent no. 4205567.9, European patent no. 0558132, 26 August 1993.
- [11] J.-C. Fort, Stabilité de l'algorithme de séparation de sources de Jutten et Herault, *Traitement Signal* 8 (1991) 35–42.
- [12] C.D. Hook, C.S. Hall, Transponder system, European patent no. 0 527 172 Bl, 5 April 1995.
- [13] A. Hyvarinen, E. Oja, A fast fixed-point algorithm for independent component analysis, *Neural Comput.* 9 (1997) 1483–1492.
- [14] C. Jutten, J. Herault, Une solution neuromimetique au probleme de separation de sources, *Traitement Signal* 5 (1988) 389–103.
- [15] C. Jutten, J. Herault, Blind separation of sources, Part I: an adaptive algorithm based on neuromimetic architecture, *Signal Process.* 24 (1991) 1-10.
- [16] O. Macchi, E. Moreau, Self-adaptive source separation, Part I: convergence analysis of a direct linear network controlled by the Herault-Jutten algorithm, *IEEE Trans. Signal Process.* 45 (1997) 918–926.
- [17] E. Moreau, Apprentissage et adaptativité. Séparation auto-adaptative de sources independantes, Ph.D. Thesis, Université Paris XI, Orsay, 1995.
- [18] E. Moreau, O. Macchi, Self-adaptive source separation, Part II: comparison of the direct, feedback and mixed linear network, *IEEE Trans. Signal Process.* 46 (1998) 39–50.
- [19] Philips Semiconductors data sheet, OM 4282 RF-Identification, Hardware description & tutorial, ID-No: 8962D26CEA20068F.
- [20] Philips Semiconductors data sheet, OM 4282 RF-Identification, Software command reference & RS 232 transmission protocol, ID-No: 7F3C1206084C995E.
- [21] E. Sorouchyari, Blind separation of sources, Part III: stability analysis, *Signal Process.* 24 (1991) 21-29.
- [22] W. Tobergte, Datenaustauschanordnung, German patent no. 4323530.1, European patent no. 0634729, 19 January 1995.

Приложение 3

Стандартизация технологии RFID

Одним из наиболее острых вопросов, препятствующих самому широкому внедрению технологии радиочастотной идентификации в мировую экономику, является стандартизация радиоинтерфейса, т.е. процесса обмена данными между считывателем и меткой при помощи электромагнитного поля. Стандартизация радиоинтерфейса способна обеспечить совместимость аппаратуры RFID, поставляемой на рынок различными производителями. Это необходимо для того, чтобы объекты с метками, произведенными в одной стране, могли бы идентифицироваться считывателями, которые произведены и установлены в других странах.

Вопросам международной стандартизации посвящена статья, подготовленная на основе материала, любезно предоставленного сотрудником компании ЕАН/ЮНИСКАН, председателем подкомитета ПК 4 «Радиочастотная идентификация» А.С. Гулария.

А.С. Гулария, С.В. Корнеев.

Современное состояние и перспективы стандартизации систем радиочастотной идентификации для управления товарными потоками в цепи поставок

Как и всякая другая высокотехнологичная отрасль человеческой деятельности, радиочастотная идентификация развивалась с момента начала своего существования как технология, применение которой ограничено рамками фирмы, которая разрабатывает оборудование для ее применения, или рамками фирмы, по заказу которой это оборудование было разработано. Начало истории развития технологии радиочастотной идентификации, как и у многих высокотехнологичных новаций нашей действительности – в недрах военно-промышленного комплекса.

1. Актуальность внедрения технологии

Как и любая другая хорошая идея, радиочастотная идентификация, в конце концов, оказалась на рынке товаров и услуг, где, как очень скоро выяснилось, её давно и с нетерпением ждали.

©А.С. Гулария, С.В. Корнеев, 2006.

Рынок, как всем известно, – система самоорганизующаяся. В недавнем прошлом самоорганизация на рынке плохо коррелировала с движения товарных масс. Движение товарной массы, как и любой другой массы, обладало избыточной инерционностью. Товары медленно продвигались и поэтому опаздывали к нервничающему потребителю или наоборот наваливались неотвратимой лавиной, что заставляло потребителя нервничать ещё больше. По этой причине рынок постоянно лихорадило, случались кризисы, связанные как с отсутствием товара, так и с его избыточным присутствием в определенном месте на рынке. У огромного количества людей, т.е. потребителей, постоянно возникали проблемы от невозможности в нужный момент получить нужный товар.

Однако избыточная инерционность движения товарных масс порождалась не столько неспособностью производителей товаров обеспечить их массовый выпуск, сколько огромным числом проблем связанных с продвижением товара от производителя к потребителю по цепи поставок. Эти проблемы являются следствием оторванности от товарного потока того потока информации, который собственно и обеспечивает перемещение товара по цепи поставки от конкретного производителя к конкретному потребителю. Без движения информации товар двигаться не может. Поэтому товарное движение сопровождается движением бумажных потоков, переносящих информацию о том, где товар произведен, куда направлен, каковы его потребительские свойства, кто товар отгрузил, кто оплатил, каким видом транспорта он перевозится, где он будет перегружаться, кто его получатель, где он пересекает границу, как он обрабатывается на таможне, сколько он находится в дороге, на каком терминале он складировается, сколько времени хранится, и, наконец, не испортился ли он. А бумага, как всем известно, не очень надежный носитель информации.

В наше время информация транспортируется по компьютерным сетям, и это существенно снижает издержки, связанные с её перемещением. Но чтобы перевести информацию с бумаги в компьютер требуется время и ручной труд. А вот времени-то у нас как раз и не хватает! И в этот напряженный для потребителя момент на рынке товаров и услуг появляется технология радиочастотной идентификации. Что же она нам предлагает? Почему же специалисты ожидают от неё столь многого? Радиочастотная идентификация наделяет товар интеллектом, она предоставляет товару возможность напрямую общаться с компьютером, а через компьютер с любым участником процесса движения товара по цепи поставок, лишая тем самым товарный поток избыточной инерционности и делая управление процессом движения товарных потоков полностью прозрачным и оптимизированным. Радиочастотная идентификация объединяет товарный и

информационный потоки в цепи поставок, сокращая тем самым огромные издержки, связанные с ручным трудом по обработке информационного потока, сопровождающего товарное движение. К чему это приведет? Результаты внедрения технологии радиочастотной идентификации в процесс управления движением товарных потоков по цепи поставок грандиозны, как по экономическому эффекту, так и по своему влиянию на развитие рынка.

2. Эффект от внедрения систем RFID

Экономический эффект от внедрения систем радиочастотной идентификации достигается за счет комплексного действия одновременно нескольких факторов, влияющих на процесс движения товаров. Во-первых, исключается ручной труд по вводу информации о конкретном товаре в компьютерную сеть, во-вторых, исключается человеческое участие в процессе организации движения от склада до витрины, в-третьих, ликвидируются ситуации, связанные с отсутствием нужного товара на витринах, в-четвертых, исключаются ситуации, связанные с нехваткой товара на складе, в-пятых, многократно возрастает информированность потребителя о наличии товара и его свойствах, в-шестых, многократно ускоряется процесс обслуживания покупателя на расчетно-кассовом узле, и, наконец, становится возможным дистанционное формирование корзины покупателя, что приводит к сокращению торговых площадей. Это ли не революция в работе с потребителем?

Столь же масштабно выглядят изменения, вызываемые внедрением систем радиочастотной идентификации, в организации процесса движения товара по цепи поставок от производителя до продавца. Автоматизация процесса ввода текущей информации о партии товара в компьютерную сеть позволяет осуществлять управление движением товара в автоматическом режиме, с одновременной оптимизацией всех временных и трудовых затрат на каждой стадии обработки товарной партии, начиная от её адресного формирования, проведения всех погрузочно-разгрузочных работ и мероприятий по её сопровождению и обработке на терминалах, до стадии пополнения склада или витрины продавца. Все стадии процесса доставки партии товара становятся прозрачными. Исключаются задержки товара в пути, связанные с потерей времени при перегрузке товарных партий, оптимизируется загрузка и маршрут движения транспортных средств, многократно ускоряется оборот возвратной тары, возрастает плавность и предсказуемость процесса поступления товаров, что позволяет сократить складские площади. Радиочастотная идентификация превращает разрозненные звенья цепи поставок в единый механизм хорошо

отлаженного конвейера для непрерывного и осмысленного продвижения товара от производителя к потребителю.

Технология радиочастотной идентификации существенным образом изменяет не только механизм распределения, но и процесс производства товаров. Использование технологии радиочастотной идентификации для маркировки деталей и продуктов производства, позволяет отслеживать в режиме реального времени все технологические стадии формирования товара в процессе его изготовления. Более того, одновременно с избирательным, осуществляемым в соответствии с требованиями конкретного заказчика, формированием конечного продукта производства – конкретного товара, формируется его технологический паспорт, который дает полное представление обо всех технологических деталях процесса изготовления этого конкретного образца товара. Не секрет, что при массовом производстве товаров, малейшее отклонение от заданного режима на любой стадии технологического цикла изготовления товара может привести к непредсказуемому изменению потребительских свойств товара. Устраняя такие технологические погрешности, фирмы-изготовители вынуждены отзываться с рынка для устранения дефектов огромные массы уже поступившего в эксплуатацию товара просто потому, что они не имеют достоверной информации о том, при изготовлении какой конкретно партии комплектующих было допущено отклонение от установленной технологии. Это незнание характеристик технологической карты каждой детали конечного продукта ведет к огромным потерям времени, труда и денег для устранения, возможно, крошечного дефекта, вызванного отступлением от заданного технологического режима. Радиочастотная идентификация, как технология формирования технологического паспорта конкретного товара, является гарантией качества его производства, особенно при современных огромных объемах производства товаров на предприятии в условиях использования комплектующих деталей от разных производителей.

Итак, формирование конкретного изделия по требованию конкретного штучного заказчика с гарантией качества в условиях массового товарного производства при снижении издержек – вот дополнительный важнейший результат внедрения технологии радиочастотной идентификации на современных предприятиях. Другим, не менее важным результатом внедрения радиочастотной идентификации является планирование загрузки производственных мощностей при работе на конкретного заказчика, а не на склад, как это всегда было в эпоху производства продукции до внедрения технологии радиочастотной идентификации.

3. Проблемы внедрения технологии

Почему же при столь привлекательных результатах использования новой технологии и на производстве, и при распределении товаров

технология радиочастотной идентификации не применяется повсеместно, на каждом заводе, на каждом складе, в каждом магазине? Не будем говорить о психологических трудностях, неизбежно возникающих, в первую очередь, у менеджеров высшего звена управления при столь значительных перестройках любого устоявшегося, отработанного годами упорного труда технологического процесса. Поговорим о реалиях, связанных с продвижением на рынок новой, глобальной по своим последствиям высокой технологии. Две основные причины, сдерживающие повсеместное применение технологии радиочастотной идентификации – это достаточно высокая стоимость одной радиочастотной метки и отсутствие глобальных стандартов.

Радиочастотная метка стоит от нескольких десятков центов до нескольких десятков долларов в зависимости от своих функциональных возможностей. Конечно, при массовом производстве стоимость её значительно снизится, но существуют сильные сомнения, что она сравняется со стоимостью бумажной этикетки со штриховым кодом, которая сегодня присутствует на каждом товаре. Поэтому сегодня радиочастотными метками маркируются объекты, стоимость которых значительно превосходит стоимость самой метки. Сегодня радиочастотная идентификация применяется для маркировки транспортных средств, возвратной тары (транспортных контейнеров, паллет, кег и других емкостей для перевозки жидкостей), контейнеров для сбора мусора и отходов, маркировки редких вин, предметов проката, например, велосипедов, книг и других бумажных носителей важной информации (в медицине это – истории болезни), инструментов и животных. Появляются первые сообщения о применении систем радиочастотной идентификации для маркировки отдельных товаров в розничной торговле и для складского учета. Так что со временем стоимость радиочастотной метки понизится настолько, что она будет присутствовать на каждой упаковке товара и на самом товаре. Но для того чтобы это произошло, производители систем радиочастотной идентификации должны придерживаться единого стандарта.

4. Стандартизация систем RFID

Поскольку производство систем радиочастотной идентификации осуществляется в разных странах, то стандарт должен быть международным, а сами системы должны быть унифицированными и совместимыми так, чтобы радиочастотные метки и устройства считывания, произведенные разными компаниями, могли безо всяких проблем работать друг с другом.

Для преодоления технических проблем, связанных с разработкой международного стандарта, крупнейшие фирмы-производители систем

радиочастотной идентификации в рамках Международной организации по стандартизации (ISO) и Международного электротехнического комитета (IEC) образовали рабочую группу, которая занимается разработкой международных стандартов систем радиочастотной идентификации, предназначенных для управления товарами. Подкомитет 31, в состав которого входит эта рабочая группа, ведет работы связанные с маркировкой товаров штриховыми кодами. Сама рабочая группа по радиочастотной идентификации делится на четыре подгруппы: Профили требований к приложениям, Синтаксис данных, Уникальная идентификация радиочастотных меток и Радиоинтерфейс. Эти подгруппы работают над созданием международных стандартов, посвященных, соответственно, общим вопросам применения систем радиочастотной идентификации, информационному наполнению радиочастотной метки и системе управления ее работой, единой системе уникальной идентификации радиочастотной метки и, наконец, правилам радиообмена, происходящего между радиочастотной меткой и устройством считывания информации. Результатом работы этих подгрупп является серия международных стандартов, полностью разрешающая все проблемы, связанные с совместимостью компонентов систем радиочастотной идентификации разных производителей.

Для упрощения процесса обеспечения необходимых функциональных возможностей систем радиочастотной идентификации стандарты определяют несколько диапазонов частот: ниже 135 кГц, 13,56 МГц, 433 МГц, 860 – 960 МГц и 2,45 ГГц. Предполагается, что системы, работающие на этих частотах, удовлетворят все потребности пользователей систем радиочастотной идентификации. Процесс разработки международного стандарта требует его согласования со всеми национальными органами стандартизации, принимающими участие в его разработке. Всего международный стандарт проходит шесть стадий согласования на разных уровнях Международной организации по стандартизации. В конце 2005 года процесс разработки международных стандартов для указанных диапазонов частот был, наконец, успешно завершен. В настоящий момент уже ведутся работы по разработке национальных стандартов технологии радиочастотной идентификации, применяемой для управления товарными потоками и предметами торговли.

В России работа по созданию стандартов радиочастотной идентификации координируется Техническим комитетом 355 «Автоматическая идентификация» Федерального агентства по техническому регулированию и метрологии. Все заинтересованные организации могут принять участие в процессе разработки российских стандартов по применению технологии радиочастотной идентификации для управления товарными потоками и предметами торговли.

Список литературы

1. www.aimglobal.org.
2. www.epcglobal.com.
3. www.gs1ru.org.

Послесловие редактора

Читатель, несомненно, может отметить, что основой этой книги является работа Т.А. Шарфельда, опубликованная в оригинале в 2001 году. К сожалению, в силу ряда причин, русская редакция этой работы появляется через 5 лет. Это огромный срок для любых новых бурно развивающихся высокотехнологичных направлений науки и техники и технология RFID тому не исключение. Тем более интересно отметить, что большая часть направлений будущих исследований, приведенных в главе 8 работы Т.А. Шарфельда, не потеряли актуальности и к настоящему времени. Так и сейчас ведутся интенсивные исследования и разработки, направленные на снижение энергопотребления и стоимости меток и считывателей, на решение многочисленных электродинамических вопросов создания систем RFID для специфических условий эксплуатации и т.д.

Как и несколько лет назад практически отсутствует серьезная научно-техническая литература, посвященная рассматриваемой технологии. Так вторым изданием в 2004 году опубликована монография Klaus Finkenzeller, RFID Handbook, а в России – достаточно интересная работа В.Л. Джхунян, В.Ф. Шаньгин. Электронная идентификация. Бесконтактные идентификаторы и смарт карты. - М.: «Издательство АСТ»: Издательство «НТ Пресс», 2004¹.

Хотя книга и не ставила перед собой целью провести обзор современного рынка, читателю, наверное, будут небезынтересны некоторые аспекты современного состояния технологии RFID.

Наиболее впечатляющий прогресс достигнут в области стандартизации радиointерфейса. В течение 2000 – 2003 гг. благодаря чрезвычайно активной работе специалистов международного Auto IDCenter, созданного на базе лабораторий Массачусеттского Технологического Института (США) и Кембриджского университета (Великобритания) были разработаны основы радиointерфейса EPC. В результате в 2006 г. появился стандарт ISO 18000-6C, который дополнил существующий стандарт ISO 18000 и, таким образом, в настоящее время в UHF диапазоне существует единый международный стандарт RFID.

Развивается рынок. Крупные мировые производственные и торговые компании Johnson&Johnson, Procter&Gamble, Coca-Cola, Wall Mart и другие финансируют реализацию крупных проектов Product Driven Supply Chain, в основе которых положены стандарты EPC и технология RFID.

¹ Наиболее свежая и полезная монография на русском языке, посвященная использованию систем RFID: Сандип Лахири **RFID. Руководство по внедрению**. Пер. с англ. – М.: КУДИЦ-ПРЕСС. – 2007.

Технологический прогресс особенно заметен в UHF диапазоне. В первую очередь это касается пассивных меток. Появление новых крупных производителей меток (Rafsec, например) свидетельствует о росте спроса с соответствующим снижением стоимости меток. Так компании-производители меток объявляют о стоимости 10 – 12 центов при объеме заказа от 1 млн. штук. Также наблюдается резкое возрастание числа производителей считывающего оборудования и специализированного программного обеспечения.

Для продвижения технологии на рынок создана всемирная организация EPC Global, которая к настоящему времени объединяет более 400 членов. Несмотря на то, что более 80% участников организации находятся в США, наибольшая динамика роста членов приходится на Европу и Юго-Восточную Азию.

Научный редактор С.В. Корнеев

Предметный указатель

Сокраще- ние	Английский термин	Русский термин	Стр.
ASK	Amplitude Shift Key	Амплитудная манипуляция	
BER	Bit Error Rate	Вероятность битовой ошибки	
BPSK	Binary Phase Shift Key	Двоичная фазовая манипуляция	
CDMA	Code Domain Multiple Access	Множественный доступ с кодовым разделением	
CEPT	European Conference of Postal and Telecommunications Administrations	Европейская организация почтовых и телекоммуникационных организаций	
CMOS	Complimentary Metal Oxide Semiconductors	Комплиментарные металлооксидные полупроводники (КМОП)	
CRC	Cycling Redundancy Checking	Циклическая избыточная проверка	
DS	Direct Sequence	Прямая последовательность	
DSP	Digital Signal Processing	Цифровая обработка сигнала	
EEPROM	Electrically Erasable Read Only Memory	Электрически перепрограммируемая память	
EIRP	Effective (or Equivalent) Isotropic Radiated Power	Эффективная (или эквивалентная) изотропная излучаемая мощность	
ERP	Effective Radiated Power	Эффективная излучаемая мощность	
ETSI	European Telecommunications Standards Institute	Европейский институт стандартов по связи	
FCC	Federal Communications Commission	Федеральная комиссия по связи США	
FDMA	Frequency Domain Multiple Access	Множественный доступ с частотным разделением	
FH	Frequency Hopping	Скачкообразная перестройка частоты	
FRAM	Ferro-electric Random Access Memory	Ферроэлектрическая память со случайным доступом	
FSK	Amplitude Shift Key	Частотная манипуляция	
ISM	Industrial, Science and Medicine (band)	Диапазон радиочастот, выделенный для использования в промышленности, науке и медицине	
MPT	Ministry of Post and Telecommunications	Министерство Почты и Телекоммуникаций Японии	
NWU	Normalized Weight Updating	Нормированный весовой алгоритм самообучения	
NRZ	Non Return to Zero	(Код) без возврата к нулю	
OOK	On Off shift Key	Переключательная манипуляция	
PPM	Phase Pulse Modulation	Фазоимпульсная модуляция (ФИМ)	

PSD	Pulse Spectral Density	Спектральная плотность мощности (СПМ)
PTM	Pulse Time Modulation	Времяимпульсная модуляция (ВИМ)
PWM	Pulse Width Modulation	Широтноимпульсная модуляция (ШИМ)
RFID	Radio Frequency Identification	Радиочастотная идентификация
ROM	Read Only Memory	Только читаемая память
RSC	Radar Cross Section	Эффективная поверхность рассеяния (ЭПР)
RTF	Reader Talk First	Считыватель излучает первым
RZ	Return to Zero	(Код) с возвратом к нулю
SNRI	Signal to Noise Ratio Improvement	Улучшение отношения сигнал/шум
SRD	Short Range Devices	Устройства малой дальности действия
TDMA	Time Domain Multiple Access	Множественный доступ с временным разделением
TTF	Tag Talk First	Метка излучает первой
UID	Unique Identification Number	Уникальный идентификационный номер
WORM	Write Once Read Many	Однократно программируемая и многократно читаемая память