

Содержание

1	Техническое задание	4
1.1	Наименование предприятия разработчика и заказчика системы	4
1.2	Основания для разработки	4
1.3	Сроки выполнения работ по созданию системы	4
1.4	Порядок оформления и предъявления результатов работы по созданию системы	5
1.5	Актуальность разрабатываемой системы	5
1.6	Краткое описание предметной области	6
1.7	Существующие аналоги	6
1.8	Назначение и цели создания системы	7
1.8.1	Система «А»	7
1.8.2	Система «Б»	9
1.8.3	Система «В»	10
1.9	Требования к системе	10
1.9.1	Требование к системе в целом	10
1.9.2	Требования к функциональным характеристикам	11
1.9.3	Требования по реализации	11
1.10	Функциональные требования к системе	12
1.10.1	Функциональные требования к системам типа «А»	12
1.10.2	Функциональные требования к системам типа «Б»	12
1.10.3	Функциональные требования к системам типа «В»	13
1.11	Входные параметры	13
1.11.1	Входные параметры систем типа «А»	13
1.11.2	Входные параметры систем типа «Б»	14
1.11.3	Входные параметры систем типа «В»	15
1.12	Требования к составу и параметрам технических средств	15
1.13	Сценарии функционирования системы	17
1.14	Состав и содержание работ по созданию системы	17
1.15	Порядок контроля и приемки системы	18
1.16	Требования к документации	18
	Список использованных источников	19

Глоссарий

Заказчик — кафедра ИУ-7 «Программное обеспечение вычислительной техники и информационные технологии» МГТУ им. Баумана.

Программное обеспечение — Совокупность программ системы обработки информации и программных документов, необходимых для эксплуатации программ;

Партнер, система-партнер, субъект РСОИ — Системы взаимодействующие в рамках разрабатываемой РСОИ.

Спам — Анонимные не запрошенные массовые рассылки электронной почты, то есть электронный эквивалент бумажной рекламной корреспонденции, засоряющей обычные почтовые ящики.

SMTP — Сетевой протокол, предназначенный для передачи электронной почты в сетях TCP/IP.

POP3 — Сетевой протокол, используемый почтовым клиентом для получения сообщений электронной почты с сервера.

MDA — Программа, принимающая входящие электронные письма и доставляющая их на электронный ящик получателя (если адрес назначения расположен на том же компьютере) или перенаправляющая их на другой почтовый сервер (если адрес назначения расположен на другом компьютере).

MTA — Программа, которая передаёт сообщения от одного компьютера к другому по электронной почте.

Обозначения и сокращения

- БД — база данных;
- ИП — интерфейс пользователя;
- МГТУ им. Баумана — Московский Государственный Технический Университет имени Н.Э. Баумана.
- ПО — программное обеспечение;
- РСОИ — распределенная система обработки информации;
- СУБД — система управления баз данных;
- ТЗ — техническое задание;
- IMAP — англ. Internet Message Access Protocol — протокол прикладного уровня для доступа к электронной почте;
- MTA — англ. Mail Transfer Agent, Message Transfer Agent — почтовый сервер;
- MDA — англ. Mail Deliver Agent — агент доставки электронной почты;
- MUA — англ. Mail User Agent — почтовый клиент;
- SMTP — англ. Simple Mail Transfer Protocol — простой протокол передачи почты;
- POP3 — англ. Post Office Protocol Version 3 — протокол почтового отделения, версия 3.

1 Техническое задание

Настоящее техническое задание разработано в рамках учебной программы по курсу «Технология программирования» на программное изделие «Распределенная система обнаружения и фильтрации спама в протоколах POP3, IMAP, SMTP». Техническое задание выполняется в соответствии со стандартом ГОСТ 34.602-89 «Техническое задание на создание автоматизированной системы».

1.1 Наименование предприятия разработчика и заказчика системы

Разработчиком системы является Сулимов Александр Сергеевич, студент группы ИУ7-104 кафедры ИУ-7 «Программное обеспечение вычислительной техники и информационные технологии» пятого курса очной формы обучения МГТУ им. Баумана.

Заказчиком системы является кафедра ИУ-7 «Программное обеспечение вычислительной техники и информационные технологии» МГТУ им. Баумана (далее Заказчик).

1.2 Основания для разработки

Разработка ведется в рамках выполнения лабораторных работ по курсу «Технология программирования» и курсового проекта по курсу «Распределенные системы обработки информации» в соответствии с учебным планом МГТУ им. Баумана на 10-й семестр для факультета ИУ кафедры ИУ-7. При выполнении работы учитываются указания, описанные в методических пособиях [1], [2].

1.3 Сроки выполнения работ по созданию системы

В соответствии с требованиями Заказчика установлены следующие сроки (указываются номера недель, начиная с 6 февраля 2012 года):

- а) начало выполнения работ — 3 неделя (20.02.2012 г.);
- б) конец выполнения работ — 15 неделя (14.05.2012 г.).

1.4 Порядок оформления и предъявления результатов работы по созданию системы

В [2] определены следующие требования к оформлению и защите проекта:

а) в процессе разработки системы необходимо использовать выделенный Заказчиком репозиторий системы контроля версий;

б) ответственное лицо со стороны Заказчика ведет контроль выполнения работы по репозиторию, учитывая количество коммитов, определяет процент выполнения проекта;

в) результаты работ представляются в форме защиты проекта, на которой необходимо предъявить:

- 1) оформленную, прошитую и подписанную руководителем проекта пояснительную записку;
- 2) презентацию в электронном виде в формате PDF;
- 3) установленное и работающее программное обеспечение с исходными кодами, хранимыми в репозитории;

г) при защите программное обеспечение должно быть установлено на компьютерах Заказчика;

д) для защиты проекта создается комиссия из двух или более ответственных лиц со стороны Заказчика.

1.5 Актуальность разрабатываемой системы

Представленные системы, позволяющие фильтровать почту от нежелательных сообщений, основываются на спам-фильтрах, которые имеют следующие недостатки:

а) необходимо обучение;

б) история обучения локальная и относится к конкретному пользователю.

Настоящая разработка должна обеспечить передачу информации между пользователями о признаках обнаруженного спама. Распределенная система обнаружения спама позволит сократить время на обучение спам-

фильтров для отдельных пользователей, тем самым повысит удобство использования почтовых клиентов.

1.6 Краткое описание предметной области

За последние десять лет сфера применения спама расширилась, а объем доставки - вырос значительно. Первое время спам рассылался напрямую на единичные адреса пользователей, и его было легко блокировать. Со временем появились высокоскоростные интернет-каналы, которые дали быструю и дешевую возможность массово рассылать спам-сообщения. Модемы пользователей не оснащались средствами защиты от несанкционированного доступа и могли использоваться злоумышленниками из любой точки планеты. Другими словами, модемы ничего не подозревающих пользователей рассылали огромные объемы спама.

Так продолжалось, пока производители аппаратного обеспечения не научились оснащать оборудование средствами защиты от спама, а спам-фильтры не стали более эффективными. Однако спам тоже эволюционировал: усовершенствовались не только способы рассылки, но и приемы, помогающие злоумышленникам обойти спам-фильтры.

В ходе анализа предметной области были рассмотрены наиболее популярные спам-фильтры и другие механизмы, применяющиеся для фильтрации потока электронной почты от спама.

1.7 Существующие аналоги

В рамках настоящей работы был произведен анализ рынка программных продуктов, позволяющих фильтровать электронную почту от спама. Было выявлено, что существует множество решений позволяющих осуществлять определение нежелательной почты для конкретного пользователя.

Существует два класса подобных продуктов:

- а) спам-фильтры на почтовых клиентах;
- б) спам-фильтры на почтовых серверах.

Спам-фильтры почтовых клиентов основаны на обработке входящих сообщений согласно заданным пользовательским правилам. Почтовые сервера используют так называемые черные и серые списки доверенных адресов входящей почты.

Таким образом, распределенных систем, позволяющих осуществлять фильтрацию спама обнаружено не было.

1.8 Назначение и цели создания системы

Назначение разработки — удовлетворить потребности клиентов (пользователей электронной почтой) в получении электронных писем, не относящихся к спаму.

Цели системы:

- а) информировать пользователя о том, что пришедшее сообщение относится к спаму;
- б) предоставлять пользователю возможность самостоятельно формировать спам-фильтры;
- в) обеспечить доставку писем, не относящихся к спаму.

Таким образом, субъектами разрабатываемой РСОИ являются:

- а) система «А», почтовый клиент (MUA);
- б) система «Б», почтовый сервер (MTA);
- в) система «С», анти-спам сервер.

Перечисленные системы в дальнейшем будут также называться субъектами РСОИ или системами-участниками.

На рисунке 1.1 представлена структура предметной области и ее составные части.

1.8.1 Система «А»

Целью систем данного типа является предоставление удобного пользовательского интерфейса для работы с электронной почтой. С точки зрения почтовой системы представляют собой MUA — программное обеспечение, устанавливаемое на компьютере пользователя и предназначенное

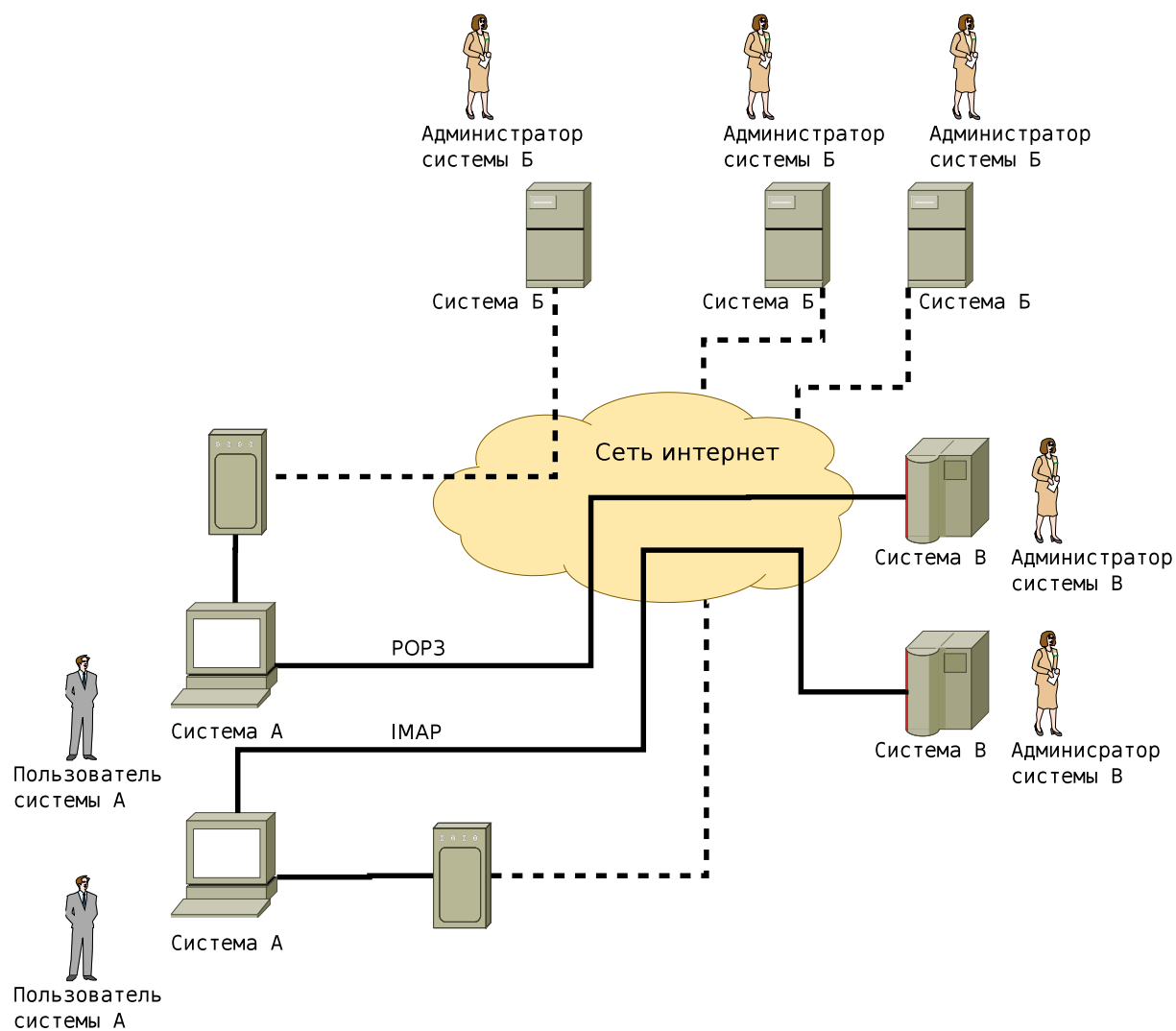


Рисунок 1.1 — Структура предметной области

для получения, написания, отправки и хранения сообщений электронной почты одного или нескольких пользователей (в случае, например, нескольких учётных записей на одном компьютере) или нескольких учётных записей одного пользователя.

Системы «А» бывают двух типов в зависимости от протокола получения входящих сообщений:

- а) POP3;
- б) IMAP.

Протокол POP3 подразумевает передачу входящих сообщений почтовым сервером и сохранение электронных писем на локальном компьютере пользователя. При использовании POP3 клиент подключается к серверу только на промежуток времени, необходимый для загрузки новых сообщений. При использовании IMAP соединение не разрывается, пока пользовательский интерфейс активен, а сообщения загружаются только по требованию клиента. Это позволяет уменьшить время отклика для пользователей, в чьих ящиках имеется много сообщений большого объёма. Протокол POP требует, чтоб текущий клиент был единственным подключенным к ящику. IMAP позволяет одновременный доступ нескольких клиентов к ящику и предоставляет клиенту возможность отслеживать изменения, вносимые другими клиентами, подключенными одновременно с ним.

1.8.2 Система «Б»

Системы аккумулируют информацию необходимую для фильтрации электронной почты пользователя от спама. Целевые пользователи систем «А» должны зарегистрироваться в одной из систем типа «Б». Такие системы получают информацию от пользователей относительно того, являются ли те или иные сообщения спамом или нет. В зависимости от полученных данных от пользователей в системах типа «Б» аккумулируется база данных, содержащая список адресов, с которых рассылаются спам-сообщения.

При получении почты системой «А» посылается запрос связанной с ней системой «Б», которая должна классифицировать входящее сообщение на наличие спама.

Системы типа «Б» взаимодействуют между собой, таким образом пользователи, входящие в состав РСОИ, обладают суммарной информацией о всех спам-рассылках в рамках текущей сети. Так как с каждым пользователем системы «А» в определенный момент времени связан лишь один антиспам сервер, то существует возможность составлять персональные спам-списки для каждого пользователя.

1.8.3 Система «В»

Системы данного типа отвечает за отправку почты и представляют собой почтовые сервера, которые обычно выполняют роль МТА и MDA. Некоторые почтовые сервера (программы) выполняют роль как МТА, так и MDA, некоторые подразумевают разделение на два независимых сервера: сервер-МТА и сервер-MDA (при этом, если для доступа к почтовому ящику используются различные протоколы — например, POP3 и IMAP, — то MDA в свою очередь может быть реализован либо как единое приложение, либо как набор приложений, каждое из которых отвечает за отдельный протокол).

1.9 Требования к системе

1.9.1 Требование к системе в целом

Разрабатываемое ПО должно удовлетворять следующим требованиям:

а) все системы в РСОИ могут быть в одном или нескольких экземплярах, включение новой системы в РСОИ не должно приводить к нарушению работы других субъектов;

б) программное обеспечение для систем каждого типа должно поддерживать функционирование системы в режиме:

- 1) системы «А» — периодическая работа в зависимости от желаний пользователя, в том числе поддерживать режим 24/7/365;

2) системы «Б» — режим 24/7/365;

3) системы «В» — режим 24/7/365.

в) системы должны быть устойчивы к отключениям питания и другим техническим сбоям, которые могут привести к нарушению нормального функционирования систем;

г) выход из строя одного субъекта РСОИ не должен приводить к сбою в работе других систем;

д) для корректного взаимодействия систем в РСОИ должен быть разработан протокол, однозначно определяющий содержание всех запросов и возможные сценарии совместной работы систем с указанием возможных состояний каждой из систем и состояний заявок в них.

1.9.2 Требования к функциональным характеристикам

К разрабатываемому ПО выдвигаются следующие функциональные требования:

а) время отклика на запрос пользователя не должно превышать 3 секунд;

б) время ожидания запросов для различных систем должно настраиваться с помощью конфигурационных файлов отдельно;

в) системы типа «Б» должны функционировать и поддерживать указанное время отклика для 50 одновременно подключенных клиентов-систем типа «А»;

г) указанные требования к временным задержкам должны соблюдаться для систем всех типов при одновременной работе с 5-10 системами партнерами.

1.9.3 Требования по реализации

При разработке РСОИ участвующие системы реализуются в виде независимых программных продуктов, использующих общий протокол прикладного уровня для взаимодействия друг с другом.

К каждому из программных продуктов предъявляются следующие требования:

- а) разрабатываемое ПО должно предоставлять удобный пользовательский интерфейс для работы администраторов и пользователей;
- б) интерфейс систем должен быть реализован как WEB-интерфейс;
- в) каждая заявка должна обладать уникальным идентификатором в рамках одного субъекта РСОИ;
- г) в качестве транспортного протокола системы должны использовать почтовые протоколы SMTP для отправки POP3/IMAP для получения сообщений и протоколы удаленного вызова процедур XML-RPC;
- д) в качестве языка разметки могут быть использованы языки XML, либо JSON;
- е) для хранения данных о пользователях, сообщениях, спаме необходимо использовать СУБД Postgres. Непосредственный доступ к базе данных одной системы должен быть закрыт для систем партнеров и клиентов.

1.10 Функциональные требования к системе

1.10.1 Функциональные требования к системам типа «А»

Система «А» должна предоставлять следующие функции:

- а) аутентификация и авторизация пользователей;
- б) просмотр входящих сообщений пользователями;
- в) регистрация анти-спам сервера через панель настроек;
- г) оповещение связанного анти-спам сервера о наличии спам сообщений;
- д) просмотр сообщений, которые являются спамом по данным РСОИ;
- е) добавление адресатов в доверенный список, после чего сообщения от данных пользователей не классифицируются как спам;
- ж) изменение способа получения входящих сообщений (протокол IMAP, POP3).

1.10.2 Функциональные требования к системам типа «Б»

Система «Б» должна предоставлять следующие функции:

- а) аутентификация и авторизация пользователей;

- б) настройка спам-фильтров на стороне почтового сервера;
- в) настройка серых списков;
- г) настройка черных списков.

1.10.3 Функциональные требования к системам типа «В»

Система «В» должна предоставить следующие функции:

- а) аутентификация и авторизация пользователей;
- б) просмотр списка подключенных почтовых клиентов к данному анти-спам серверу;
- в) просмотр списка адресатов, признанных за спамеров;
- г) настройка списка адресатов;
- д) просмотр статуса входящего сообщения;
- е) отмена обработки сообщения: доставка клиенту, удаление.

1.11 Входные параметры

1.11.1 Входные параметры систем типа «А»

Система должна содержать следующую информацию о анти-спам сервере (система типа «В»):

- а) идентификатор системы;
- б) название системы;
- в) веб-адрес системы;
- г) адрес электронной почты системы;
- д) адрес XML-RPC сервера системы.

Система должна содержать следующую информацию о почтовом сервере (система типа «Б»):

- а) идентификатор;
- б) название системы;
- в) веб-адрес системы;
- г) адрес электронной почты системы;
- д) адрес исходящей почты;
- е) адрес входящей почты и протокол ее получения.

Система должна содержать следующую информацию о прокси-серверах POP3 и IMAP:

- а) веб-адрес сервера.

Система должна содержать следующую информацию о пользователе:

- а) логин(уникальный идентификатор, адрес электронной почты);
- б) пароль;
- в) время последнего входа в систему;
- г) способ работы в настоящий момент времени (автономно, либо удаленно).

Система должна содержать следующую информацию о входящем сообщении:

- а) уникальный идентификатор сообщения в рамках системы;
- б) наименование отправителя/получателя;
- в) контент сообщения (тема и тело сообщения);
- г) маркер спама, установленный пользователем (оценка от 0% до 100%);
- д) время получения сообщения.

Система должна содержать следующую информацию об ошибках в работе:

- а) уникальный идентификатор ошибки в рамках системы;
- б) время возникновения;
- в) код ошибки;
- г) логин пользователя;
- д) адрес почтового сервера, адрес анти-спам сервера, адреса прокси серверов.

1.11.2 Входные параметры систем типа «Б»

Система должна содержать следующую информацию о сообщении:

- а) уникальный идентификатор сообщения в рамках системы;

- б) адрес отправителя;
- в) адрес получателя;
- г) контент сообщения.

Система должна знать следующую информацию об администраторах системы:

- а) логин;
- б) пароль.

Система должна содержать следующую информацию об ошибках в работе:

- а) уникальный идентификатор ошибки в рамках системы;
- б) время возникновения;
- в) код ошибки.

1.11.3 Входные параметры систем типа «В»

Система должна знать следующую информацию о сообщении:

- а) уникальный идентификатор сообщения в рамках системы;
- б) адрес отправителя;
- в) адрес получателя;
- г) контент сообщения.

Система должна содержать следующую информацию об ошибках в работе:

- а) уникальный идентификатор ошибки в рамках системы;
- б) время возникновения;
- в) код ошибки;
- г) статус сообщений.

1.12 Требования к составу и параметрам технических средств

Минимальные требования к программно-аппаратному обеспечению для рабочей станции системы типа «А»:

- а) тактовая частота — не менее 3,0 ГГц;
- б) оперативная память — не менее 1 Гб;
- в) требования к операционной системе:
 - 1) Windows XP и старше;
 - 2) Mac OS X 10.5 и старше;
- г) свободное пространство на жестком диске не менее 10 Гб для операционной системы;
- д) свободное пространство на жестком диске не менее 200 Мб для программного обеспечения;
- е) Ethernet-адаптер стандарта 1000BASE-T для доступа к сети;
- ж) веб-браузер;
- з) интерпретатор языка Python (для работы прокси).

Минимальные требования к программно-аппаратному обеспечению для почтовых серверов систем типа «Б»:

- а) тактовая частота — не менее $2 \times 3,0$ ГГц;
- б) оперативная память — не менее 4 Гб;
- в) требования к операционной системе:
 - 1) Windows Server 2003 и старше;
 - 2) дистрибутив линукс с ядром версии 2.6:3.0;
- г) свободное пространство на жестком диске не менее 10 Гб для операционной системы;
- д) свободное пространство на жестком диске не менее 200 Мб для программного обеспечения;
- е) Ethernet-адаптер стандарта 1000BASE-T для доступа к сети;
- ж) интерпретатор языка Python (для работы почтового сервиса);
- з) СУБД Postgres;
- и) платформа Django для веб-интерфейса.

Минимальные требования к программно-аппаратному обеспечению для анти-спам серверов систем типа «В»:

- а) тактовая частота — не менее $2 \times 3,0$ ГГц;
- б) оперативная память — не менее 4 Гб;

- в) требования к операционной системе:
 - 1) Windows Server 2003 и старше;
 - 2) дистрибутив линукс с ядром версии 2.6:3.0;
- г) свободное пространство на жестком диске не менее 10 Гб для операционной системы;
- д) свободное пространство на жестком диске не менее 200 Мб для программного обеспечения;
- е) Ethernet-адаптер стандарта 1000BASE-T для доступа к сети;
- ж) интерпретатор языка Python (для работы прокси);
- з) платформа Django для веб-интерфейса;
- и) СУБД Postgres.

1.13 Сценарии функционирования системы

1.14 Состав и содержание работ по созданию системы

В таблице 1.1 представлен состав работ по созданию РСОИ «Распределенная система обнаружения и фильтрации спама в протоколах POP3, IMAP, SMTP». Перечень работ соответствует требованиям Заказчика.

Таблица 1.1 — Перечень работ по созданию РСОИ

Выполняемая работа	Срок выполнения
Исследование объектов автоматизации, сбор сведений о существующих аналогах	1 неделя
Разработка технического задания	4 неделя
Разработка пилотного проекта по выбранному варианту РСОИ	6 неделя
Разработка окончательных решений по выбранным структурам, разработка конечных вариантов процедур и заглушек систем-партнеров	8 неделя
Разработка пользовательского интерфейса	12 неделя
Создание документации	12 неделя
Отладка проекта, подготовка к защите	14 неделя
Защита проекта	15 неделя

1.15 Порядок контроля и приемки системы

При разработке системы необходимо произвести следующие испытания:

а) тестирование логики работы систем всех типов «А», «Б», «В»: при тестировании системы одного типа системы другого типа представляются в виде заглушек, работающих по строго заданным сценариям;

б) тестирование нормальной совместной работы систем «А», «Б», «В»: проверяются различные сценарии работы при корректных входных данных, предполагаемые результаты работы сравниваются с реально полученными;

в) испытание РСОИ на отказоустойчивость: имитирование таких событий как отключение питания, выход из строя других систем в составе РСОИ, поступление запросов с ошибочной или противоречивой информацией, поступление запросов в неверном формате и порядке.

1.16 Требования к документации

Разработка, установка и внедрение системы должны сопровождаться следующими документами:

а) руководство по установке, настройке систем всех типов при развертывании РСОИ;

б) руководство по использованию систем «А» для пользователей;

в) руководство по использованию систем «Б», «В» для администраторов.

Список использованных источников

1. *Романова, Т.Н.* Методическое пособие / Т.Н. Романова. — М., 2009.
2. *Крищенко, В.А.* Методическое пособие / В.А. Крищенко. — М., 2009.