Б. В. Черников

ЗАЩИТА ИНФОРМАЦИИ ПРИ ПЕРЕДАЧЕ СЛАБОФОРМАЛИЗУЕМЫХ ДОКУМЕНТОВ

В связи с широким распространением персональных компьютеров не только как средств обработки информации, но и как оперативных средств коммуникации (электронная почта), возникают проблемы, связанные с обеспечением защиты информации от ее перехвата, преднамеренных или случайных искажений. Развитие информационных технологий сопровождается, к сожалению, ростом компьютерных преступлений, связанных с хищением конфиденциальной и другой информации, а также обусловленных этим обстоятельством материальных потерь. Первое компьютерное преступление, совершенное в городе Миннеаполисе в 1958 г., состояло в подделке банковских документов с помощью компьютера. По некоторым данным, утечка 20 % коммерческой информации в 60 % случаев приводит к банкротству фирмы. И это немудрено, поскольку по существующей статистике при ограблении банка потери (в среднем) составляют 19 тысяч долларов, а при компьютерном преступлении — 560 тысяч долларов.

Актуальность проблемы защиты информации подчеркивается тем обстоятельством, что персональный компьютер или автоматизированное рабочее место является частью систем обработки информации, систем коллективного пользования, вычислительных сетей. Причины активизации компьютерных преступлений заключаются именно в том, что информационные технологии позволили реализовать идею академика В. М. Глушкова о безбумажных технологиях [1], создающих «…прочную основу для перестройки управления социально-экономическими процессами на основе безбумажной технологии в масштабах целой отрасли, крупного региона и даже целой страны». Однако обоснованное им еще в XX в. объединение вычислительных машин в крупные сети, впоследствии реализованное в виде всемирной сети Интернет, вызвало необходимость в предъявлении достаточно жестких требований к надежности и достоверности передаваемой информации, к предотвращению несанкционированного доступа к документам, передаваемым по сетям связи.

Юридическая значимость информации приобретает важность в последнее время. Одной из причин этого является создание и развитие нормативно-правовой базы безопасности информации в масштабах государства. Вопросы защиты информации рассматриваются на государственном уровне, являясь предметом федеральных законов Российской Федерации [1, 2, 3, 4].

Федеральный закон № 149–ФЗ «Об информации, информационных технологиях и о защите информации» рассматривает защиту информации как «…принятие правовых, организационных и технических мер, направленных на обеспечение защиты информации от неправомерного доступа, уничтожения, модифицирования, блокирования, копирования, предоставления, распространения, а также от иных неправомерных действий в отношении такой информации», и, кроме того, мер, направленных на «…соблюдение конфиденциальности информации ограниченного доступа».

Столь внимательный подход государственных органов власти к указанным вопросам свидетельствует об актуальности и значимости проблемы, причем в соответствии со ст. 6 Федерального закона № 149–ФЗ обязанность по защите информации возлагается на обладателя информации, что доказывает необходимость совершенствования мер защиты информации в различных сферах деятельности.

Организация защиты информации

Применение мер защиты к информации необходимо, поскольку в конечном счете она в дальнейшем материализуется в продукцию или услуги, приносящие прибыль. При недостаточном уровне защиты информации резко возрастает вероятность снижения прибыли и появления убытков вследствие вторжения злоумышленников в информационное пространство организации (Рис. 1).

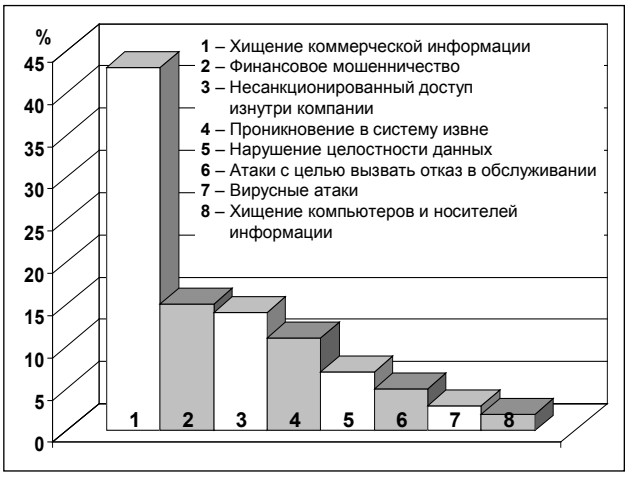

Рис. 1. Частота осуществления атак на информационные системы

Из данных, приведенных на рис. 1, видно, что почти четверть атак на информационные системы напрямую обусловлены несанкционированным доступом к информации, что является весьма значимым фактором, определяющим необходимость использования и постоянного совершенствования средств защиты.

Традиционными способами разграничения доступа к конфиденциальной информации изначально являлись организационные меры, основанные лишь на соблюдении сотрудниками процедуры допуска к информации. Организационные меры эффективно обеспечивали разграничение доступа к информации при организации работы на основе твердых (бумажных) копий документов, однако с развитием компьютерных систем эти меры перестали обеспечивать необходимую безопасность информации.

Информация, создаваемая и распространяемая с помощью средств вычислительной техники, может быть изменена, к ней может быть получен несанкционированный доступ. С позиций требований информационной безопасности наиболее опасным считается именно несанкционированный доступ.

В настоящее время широко применяются специализированные программные и программно-аппаратные средства защиты информации, которые позволяют максимально автоматизировать процедуры доступа к информации и обеспечить при этом требуемую степень ее защиты.

Система защиты информации, являясь неотъемлемой составной частью системной архитектуры, строится в соответствии с моделью, которая формируется на основе требований нормативных документов государства в области защиты информации, а также адаптации международных стандартов информационной безопасности в условиях действующего нормативно-правового поля РФ. К нормативным документам прежде всего необходимо отнести руководящие документы Федеральной службы по техническому и экспортному контролю, имеющие первостепенное значение в нашей стране [1, 2, 3, 4, 5].

Достаточно полно критерии для оценки механизмов безопасности организационного уровня представлены в международных стандартах ISO/IEC 17799:2005 и ISO/IEC 27001:2005, принятых в 2005 г. Они содержат практические правила по управлению информационной безопасностью и могут использоваться в качестве критериев для оценки механизмов безопасности организационного уровня, включая административные, процедурные и физические меры защиты [1, 2]. Критерии для оценки механизмов безопасности программно-технического уровня представлены в международном стандарте ISO/IEC 15408:2002, принятом в 2002 г. Этот стандарт определяет функциональные требования безопасности (Security Functional Requirements) и требования к адекватности реализации функций безопасности (Security Assurance Requirements) [1].

Организационные меры не в состоянии предотвратить в полной мере попытки несанкционированного доступа, поскольку они распространяются исключительно на масштабы организации, не охватывая каналы связи, и не предполагают применения технических средств борьбы с угрозами перехвата информационных сообщений. Наряду с применением разных приоритетных режимов и систем разграничения доступа разработчики информационных систем уделяют внимание различным криптографическим методам обработки информации.

Большинство средств криптографической защиты данных реализуется с помощью специализированных *аппаратных* устройств, устанавливаемых на передающей и приемной сторонах, — шифратора и дешифратора, которые осуществляют соответственно шифрование и дешифрование передаваемой информации. Применение специализированной аппаратуры для шифрования обусловливает относительно высокую стоимость реализации, однако наблюдается определенное преобладание аппаратных методов по сравнению с программными. В основном рассматриваются преимущества аппаратных решений, относящиеся к скорости обработки информации и обеспечению физической защиты компонентов. Считается, что аппаратные средства в состоянии более быстро осуществить необходимые операции по обработке данных, чем программы могут реализовать сложные криптографические алгоритмы.

Программное шифрование представляет собой результат реализации криптографического алгоритма программными средствами. Достоинства в использовании программных средств заключаются в возможности тиражирования путем обычного копирования, в относительной простоте модификации и использования. Кодирование текстовой информации может проводиться фактически с помощью кодовых таблиц путем замены одних символов другими. При этом может осуществляться и определенное сжатие передаваемого информационного пакета. Если информация зашифрована с помощью простой подстановки, то расшифровать ее можно было бы, определив частоты появления каждой буквы в шифрованном тексте и сравнив их с частотами букв русского алфавита. Таким образом, существует возможность определения подстановочного алфавита и расшифровывается текст.

Анализ методов шифрования, применяемых в настоящее время, показывает, что, несмотря на достаточно широкое их использование, они не вполне свободны от недостатков и оставляют определенное поле для совершенствования и разработки новых методов защиты информации, передаваемой по каналам связи.

Защита слабоформализуемых документов на основе лексикологического синтеза

Документ представляет собой сложную информационную совокупность, характеризующуюся множеством различных параметров (состав реквизитов, их содержание, формат, тип носителя, правила расположения информации по полю документа и т. д.), каждый из которых может быть принят за объект унификации. Несмотря на огромное число разнотипных документов с множеством параметров, в каждом из них различают форму и содержание.

Используемые методы шифрования информации не обеспечивают передачи данных в виде документа определенного типа, имеющего конкретное расположение реквизитов в соответствии с установленной формой данного документа. Значимым компонентом является лишь содержательная часть документа. В то же время документ целесообразно восстанавливать не только в аспекте содержания, но и сохраняя его форму. Тем самым, наряду со снижением вероятности несанкционированного ознакомления с текстом документа при его передаче по каналам связи и сокращением стоимости шифрования и расшифровки документа, будет сохранена не только содержательная часть документа, но и его форма с учетом расположения его реквизитов.

Слабоформализуемые документы — полнотекстовые документы, содержание которых существенным образом связано с произвольной, меняющейся в зависимости от конкретной ситуации структурой. Это документы, обладающие достаточно высокой степенью вариативности. В связи с этим содержательная структуризация слабоформализуемых документов может требовать

детализации как взаимосвязи, так и взаимной зависимости композиции текста вплоть до атомарных значений — фрагментов фраз, слов и даже частей отдельных слов.

Перспективным в плане решения задачи дополнительной защиты документов при передаче их по каналам связи представляется использование возможностей автоматизированного способа лексикологического синтеза документов [1].

Поставленная задача может быть решена посредством передачи специализированной информационной посылки, формируемой в процессе лексикологического синтеза документа, с последующим ее восстановлением на приемной (адресной) стороне. Эта информационная посылка должна содержать индексную последовательность, определяющую номер формы документа и условные номера опорных слов, выбранных по лексикологическому дереву сформированного документа, с дополнением введенной неунифицированной информации. На адресной стороне должно проводиться автоматизированное восстановление сформированного документа на основе полученной индексной последовательности с помощью программы, аналогичной той, с применением которой формировался первоначальный документ. Для этого следует последовательно пройти по установленным ветвям лексикологического дерева документа, восстанавливая необходимые опорные слова для внедрения формулировок, относящихся к категории переменной унифицированной информации, в документ, формируемый на приемной стороне.

Сущность предлагаемого способа защиты информации при передаче документов по каналам связи [1] приведена на рис. 2.

Рис. 2. Сущность защиты информации на основе лексикологического синтеза документов

В соответствии с задачами, решаемыми в определенной сфере деятельности, устанавливается совокупность реквизитов каждого документа, расположенных в определенной последовательности, и определяется местоположение в документе каждого элемента информации. Это позволяет

разработать формуляры документов или формуляры-образцы (унифицировать форму документа). Таким образом определяется модель построения документа. При этом устанавливаются основной набор реквизитов официального письменного документа, размеры полей, требования к формату документа. Определяются зоны унифицированной формы документа, предназначенные для закрепления ее в технических средствах хранения документов, а также для нанесения специальных изображений. Определяется набор реквизитов, необходимых и достаточных для идентификации автора официального документа, которые в дальнейшем в целях экономии времени будут постоянно вноситься в бланк документа программными средствами. Перечисленные компоненты документа определяют его форму. Совокупность форм собирается в базу форм документов, причем каждая форма должна иметь свой индекс, который выбирается на стадии подготовки передаваемого документа и фиксируется для последующей передачи совместно с последующей индексной последовательностью пройденного маршрута по лексикологическому дереву документа.

Индексная последовательность, соответствующая формируемому документу, фиксируется в процессе его создания при прохождении лексикологического дерева. Сформированная информационная посылка передается по каналам связи и на приемной стороне является исходной для восстановления документа.

По индексу формы восстанавливается унифицированная форма документа, оставшаяся часть принятой индексной последовательности используется для восстановления содержательной части документа.

На рис. 3 приведена структура процесса автоматизированной фиксации индексной последовательности, соответствующей содержательной части документа, при его формировании на передающей стороне.

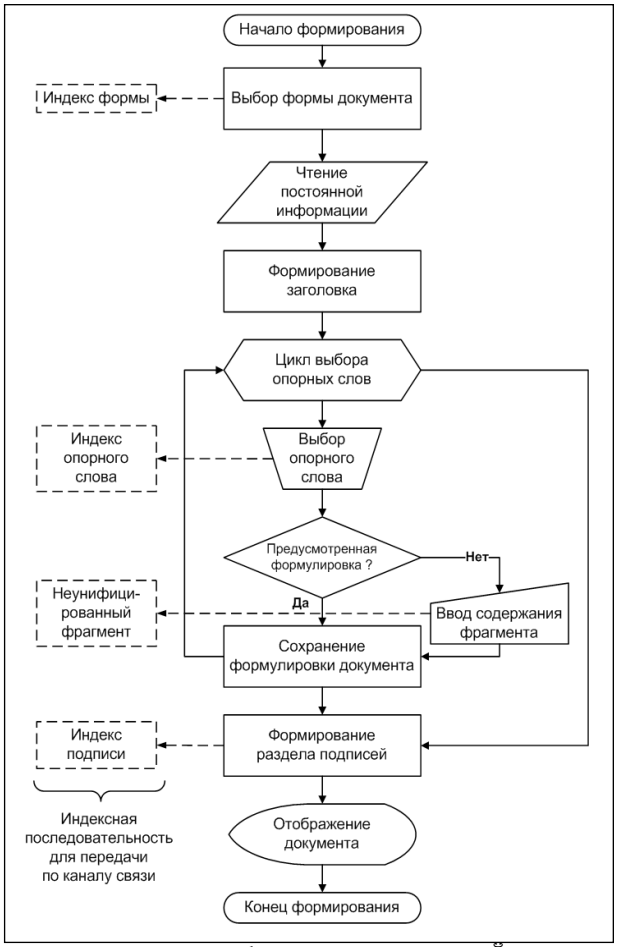

Рис. 3. Структура процесса фиксации индексной последовательности

Постоянная информация, чтение которой предусмотрено при формировании заголовка программными средствами, может содержать сведения о наименовании организации, структурном подразделении, набор реквизитов организации, наименование документа и другие подобные данные, редко подверженные изменениям и содержащиеся в отдельной структуре сохраняемой информации.

Далее организуется цикл выбора опорных слов по лексикологическому дереву, предусмотренному для формирования документа выбранного типа. В рамках этого цикла при формировании документа выбирается очередное опорное слово, индекс которого фиксируется. Если выбранному опорному слову соответствует предусмотренная формулировка, то она внедряется в документ. В случае отсутствия предусмотренной унифицированной формулировки в документ, как и в информационную посылку, будет внедрена формулировка, относящаяся с категории переменной неунифицированной информации.

По завершении цикла опорных слов формируется раздел подписей документа, при этом в индексную последовательность внедряются соответствующие индексы подписей сотрудников, визирующих формируемый документ.

Сформированный документ отображается на экране монитора для контроля, а сформированная индексная последовательность подготовлена для передачи по каналам связи.

На рис. 4 приведена лексикологическая схема примера фиксации фрагмента индексной последовательности для передачи протокола осмотра пациента медицинского учреждения при проведении гастроскопии[1].

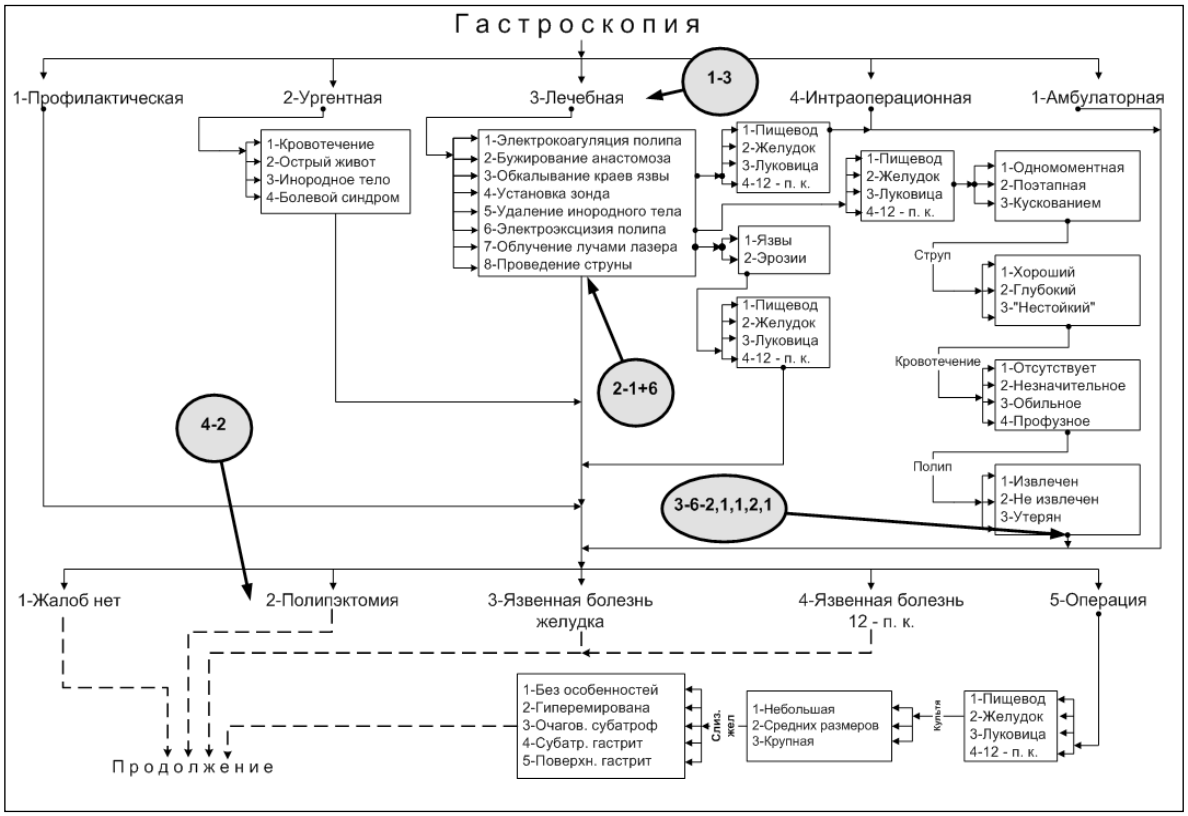

Рис. 4. Пример лексикологической схемы протокола гастроскопии при фиксации индексной последовательности

Выбор того или иного опорного слова означает выбор требуемого индекса компонента в лексикологическом дереве документа. При выборе опорных слов индекс выбираемых слов фиксируется, составляя в совокупности документарную индексную последовательность,

[1] Пример приводится для случая применения рассматриваемого способа в сфере медицины ввиду чрезвычайно высокой вариативности формируемых в этом направлении деятельности документов и, следовательно, наибольшей демонстрации эффективности.

соответствующую пройденным опорным пунктам маршрута по лексикологическому дереву в соответствии со следующими уровнями уточнения информации:

— первый уровень — уровень типа гастроскопии;

— второй уровень — уточнение типа (причины ургентной гастроскопии, манипуляции при лечебной гастроскопии);

— третий уровень — конкретизация манипуляций при лечебной гастроскопии;

— четвертый уровень — характеристика состояния пациента;

— пятый уровень — характеристика оперативных действий.

На лексикологической схеме, например, показано, что при выборе типа гастроскопии можно выбрать лечебную. В этом случае для уровня типа гастроскопии 1 фиксируется индекс 3.

При выборе лечебной гастроскопии необходимо далее выбирать нужные манипуляции. Пусть пользователем выбраны манипуляции «Электрокоагуляция полипа» и «Электроэксцизия полипа». В этом случае для уровня 2 фиксируется индексная последовательность «1+6», которая означает совместный выбор позиций 1 и 6.

В третьем уровне для позиции 6 второго уровня необходимы конкретизирующие позиции. Пусть выбрана одномоментная электроэксцизия полипа для отдела «Желудок» с хорошим струпом, незначительным кровотечением и извлечением полипа. В этом случае фиксируется индексная последовательность «3–6–2,1,1,2,1».

На четвертом уровне пусть выбирается характеристика «Полипэктомия», в этом случае фиксируется индекс «4–2».

В целом для документального описания зафиксированная индексная последовательность, включаемая в документарный индексный пакет, может выглядеть следующим образом:

1–1+2+4;2–3;3–0;4–1+2+3;6–1;7–1;8–1+3+4;9–2+4;10–1+5+7.

С учетом содержательной части формируемого документа, а также текстовых элементов, присущих данному виду документа, приведенный пример может быть аналогом фрагмента (в документе не использованы графические компоненты), приведенного на рис. 5. Даже без подробного анализа достаточно хорошо виден высокий уровень защиты информации, содержащейся в документе, восстановить который без наличия исходной модели (лексикологического дерева) не представляется возможным.

Отделение эндоскопии и гастроэнтерологических исследований

Протокол № G/3-2/1 от 15 января 2009 г.

ПЕРВИЧНАЯ АМБУЛАТОРНАЯ ЛЕЧЕБНАЯ
ЭЗОФАГОГАСТРОДУОДЕНОФИБРОСКОПИЯ

ФАМИЛИЯ Имя Отчество Пациента
Год рождения – 1950 Пол – мужской История болезни № 12345
4 отделение

Исследование проводилось аппаратом G-3. Проведена анестезия Sol.Lidocaini 10% – Spray. Жалобы: Аппарат введен в пищевод, желудок, луковицу 12-перстной кишки.
Слизистая дистальной трети пищевода гиперемирована. Кардия смыкается не полностью. Отмечается пролабирование слизистой желудка в пищевод. В желудке повышенное количество содержимого. Содержимое желудка окрашено желчью. Слизистая желудка ярко гиперемирована, с картиной субатрофического гастрита. Угол желудка без особенностей. Антральный отдел формируется правильно. Пилорус проходим, спазмирован, гиперемирован.

Рис. 5. Пример фрагмента содержательной части документа

Сформированная индексная последовательность передается по каналу связи. Как видно из содержания передаваемой индексной последовательности, восстановление текста сформированного документа невозможно без специализированной обработки, поскольку сама индексная последовательность не содержит данных, по которым можно было бы воссоздать исходный текст путем каких-либо операций по перекодированию.

На приемной стороне осуществляется восстановление первичного документа. На рис. 6 приведена блок-схема последовательности операций, иллюстрирующая процесс автоматизированного восстановления документа на приемной стороне на основе зафиксированной индексной последовательности.

На приемной стороне в автоматизированном режиме осуществляется лексикологический синтез, т. е. формируются текстовые фрагменты с помощью компьютерной системы путем создания фраз на основе использования набора опорных (ключевых) слов, комплектуемого в соответствии с содержанием переданной индексной последовательности. Наряду с синтезом текстовых фрагментов документа проводится автоматическое связывание фрагментов и отдельных слов текста в соответствии с правилами орфографии и лексикологии. Необходимую связь между словами в используемых фразах путем некоторого изменения отдельных слов в формулировках в целях их согласованного применения (с точки зрения правил синтаксиса) обеспечивают программные средства.

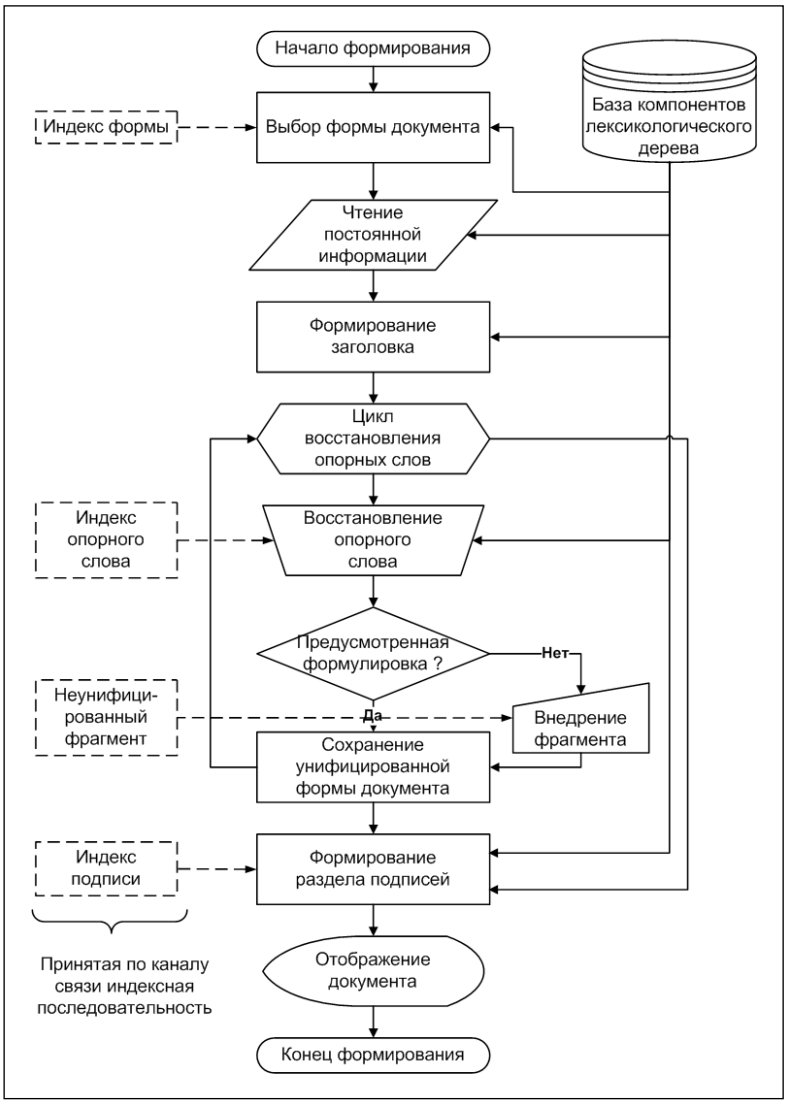

Рис. 6. Блок-схема процесса автоматизированного восстановления документа на приемной стороне

Автоматизированное формирование исходного документа осуществляется с использованием специализированной программы и стандартного компьютера. Формирование документа ведется в диалоговом режиме с автоматическим пошаговым «наращиванием» объема текста за счет внедрения

конкретных формулировок, связанных с зафиксированными индексами опорными словами. Унифицированная постоянная информация внедряется в документ автоматически.

Восстановление формы документа осуществляется из базы компонентов лексикологического дерева на основе индекса формы, после чего постоянная информация считывается из базы данных и формируется заголовок документа.

Организуется цикл восстановления опорных слов для имитации прохождения по лексикологическому дереву формируемого документа. В рамках этого цикла на основе принятых индексов опорных слов считываются из базы данных требуемые опорные слова. Затем производится проверка, предусмотрена ли в числе сохраненных формулировка, относящаяся к выбранному опорному слову. Если формулировка не предусмотрена (т. е. ответ на вопрос «НЕТ»), то внедряется содержание свободной формулировки. Затем в текст формируемого документа добавляется очередная формулировка. Далее производится переход к следующему шагу цикла, который повторяется до исчерпания индексов опорных слов лексикологического дерева документа в принятой индексной последовательности. Затем формируется раздел подписей документа, для которого используют принятый индекс подписей и считанную в соответствии с ним из базы данных информацию. Для проверки выполненной работы предусмотрено отображение сформированного документа на экране монитора.

Структура системы, реализующей способ защиты информации при передаче документов по каналам связи, приведена на рис. 7.

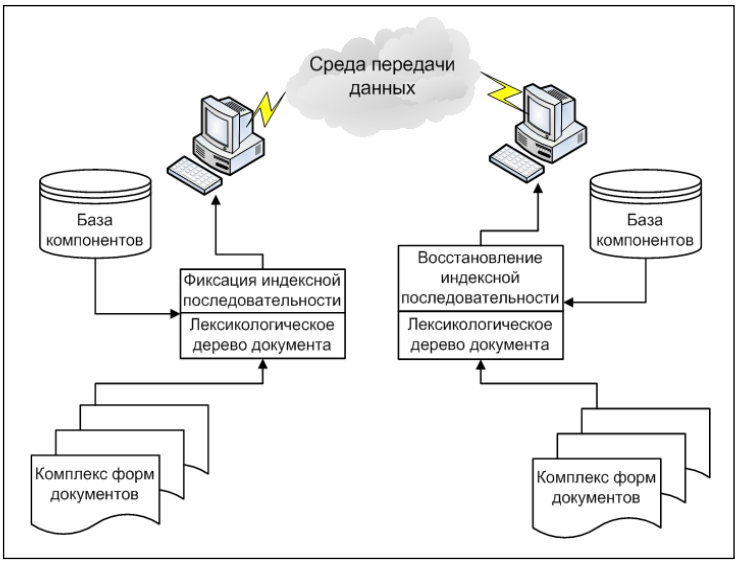

Рис. 7. Структура системы, реализующей способ защиты информации при передаче документов по каналам связи

На передающей стороне с помощью лексикологического дерева документа, связанного с базой данных и комплексом форм документов, фиксируется индексная последовательность формируемой информации, которая передается по каналам связи через среду передачи данных. На приемной стороне осуществляется восстановление индексной последовательности при использовании согласованного лексикологического дерева документа, связанного с таким же комплексом форм документов и базой данных, содержащей заготовки фрагментов документа, формируемого путем прохождения по лексикологическому дереву.

Анализ результатов экспериментальной проверки предлагаемого способа защиты информации при передаче документов по каналам связи показывает практическую невозможность несанкционированного восстановления документов при их передаче по каналам связи на основе

зафиксированной индексной последовательности при отсутствии согласованного лексикологического дерева на передающей и приемной сторонах. Дополнительным достоинством является возможность восстановления не только содержания, но и формы передаваемого документа.

СПИСОК ЛИТЕРАТУРЫ:

1. *Глушков В. М.* Основы безбумажной информатики. М.: Наука, 1984.
2. Федеральный закон Российской Федерации от 21 июля 1993 г. № 5485–1 (ред. от 01.12.2007) «О государственной тайне».
3. Федеральный закон Российской Федерации от 29 июля 2004 г. № 98–ФЗ «О коммерческой тайне».
4. Федеральный закон Российской Федерации от 27 июля 2006 г. № 149–ФЗ «Об информации, информационных технологиях и о защите информации».
5. Федеральный закон Российской Федерации от 27 июля 2006 г. № 152–ФЗ «О персональных данных».
6. Руководящий документ ФСТЭК от 30 марта 1992 г. «Концепция защиты средств вычислительной техники и автоматизированных систем от несанкционированного доступа к информации».
7. Руководящий документ ФСТЭК от 30 марта 1992 г. «Защита от несанкционированного доступа к информации. Термины и определения».
8. Руководящий документ ФСТЭК от 30 марта 1992 г. «Классификация автоматизированных систем и требования по защите информации».
9. Руководящий документ ФСТЭК от 30 марта 1992 г. «Временное положение по организации разработки, изготовления и эксплуатации программных и технических средств защиты информации от несанкционированного доступа в автоматизированных системах и средствах вычислительной техники».
10. Руководящий документ ФСТЭК от 25 июля 1997 г. «Защита от несанкционированного доступа к информации. Показатели защищенности от несанкционированного доступа к информации».
11. Международный стандарт ISO/IEC 17799:2005 «Information technology — Security techniques — Code of practice for information security management»
12. Международный стандарт ISO/IEC 27001:2005 «Information technology — Security techniques — Information security management systems — Requirements».
13. Международный стандарт ISO/IEC 15408:2002 «Common Criteria for Information Technology Security Evaluation».
14. *Черников Б. В.* Способ автоматизированного лексикологического синтеза документов. Патент РФ №2253893, 2005.
15. *Черников Б. В.* Способ автоматизированного формирования документов с защищенной информацией при передаче их по каналам связи. Патент РФ №2331104, 2008.