1.1. Основные термины и определения

Основные положения по обеспечению защиты информации приведены в следующих документах:

- Закон Российской Федерации от 06.04.11 № 63-ФЗ «Об электронной подписи».
- Закон Российской Федерации от 04.05.11 № 99-ФЗ «О лицензировании отдельных видов деятельности»
- Закон Российской Федерации от 27.07.06 № 149-ФЗ «Об информации, информационных технологиях и о защите информации».
- Закон Российской Федерации от 07.07.03 № 126-ФЗ «О связи».
- Закон Российской Федерации от 21.07.93 № 5485-1 «О Государственной тайне».
- Закон Российской Федерации от 29.06.15 № 162-ФЗ «О стандартизации в Российской Федерации».
- Закон Российской Федерации от 17.01.97 № 85-ФЗ «Об участии в международном информационном обмене».
- Закон Российской Федерации от 27.07.06 № 152-ФЗ «О персональных данных».
- Приказ ФАПСИ от 13.06.01 № 152 «Об утверждении инструкции об организации и обеспечении безопасности хранения, обработки и передачи по каналам связи с использованием средств криптографической защиты информации с ограниченным доступом, не содержащей сведений, составляющих государственную тайну».
- Гражданский кодекс Российской Федерации.

Основные термины и определения сформулированы в следующих стандартах:

- ГОСТ Р 51275-2006 «Защита информации. Объект информатизации. Факторы, воздействующие на информацию. Общие положения».

- ГОСТ Р 50.1.056-2005 «Техническая защита информации. Основные термины и определения».
- ГОСТ Р 50.1.053-2005 «Информационная технология. Основные термины и определения в области технической защиты информации».

Основные термины, относящиеся к обеспечению защиты информации (в алфавитном порядке):

- *spyware* – программа, которая скрытным образом устанавливается на компьютер с целью сбора информации о конфигурации компьютера, пользователе, пользовательской активности или выполнения иных действий без согласия последнего.
- *антивирусная программа* – программа, предназначенная для поиска, обнаружения, классификации и удаления компьютерного вируса и вирусоподобных программ.
- *безопасность автоматизированной информационной системы* – состояние защищенности автоматизированной информационной системы, при котором обеспечиваются конфиденциальность, доступность, целостность, подотчетность и подлинность ее ресурсов.
- *безопасность информации* – состояние защищенности информации, при котором обеспечиваются ее конфиденциальность, доступность и целостность (безопасность информации определяется отсутствием недопустимого риска, связанного с утечкой информации по техническим каналам, несанкционированными и непреднамеренными воздействиями на данные и (или) на другие ресурсы автоматизированной информационной системы, используемые при применении информационной технологии).
- *безопасность информации (при применении информационных технологий)* – состояние защищенности информационной технологии, обеспечивающее безопасность информации, для обработки которой она применяется, и информационную безопасность автоматизированной информационной системы, в которой она реализована.

- *ботнет* – это компьютерная сеть, состоящая из некоторого количества хостов, с запущенными ботами – автономным программным обеспечением. Чаще всего *бот* в составе ботнета является программой, скрытно устанавливаемой на устройство жертвы и позволяющей злоумышленнику выполнять некие действия с использованием ресурсов заражённого компьютера (рассылки спама, перебора паролей на удалённой системе, атак на отказ в обслуживании и так далее).
- *бэкдор* – программа, которая устанавливается взломщиками на компьютере после получения первоначального доступа с целью повторного получения доступа к системе.
- *вредоносная программа* – программа, предназначенная для осуществления несанкционированного доступа и (или) воздействия на информацию или ресурсы автоматизированной информационной системы.
- *доступность информации* – состояние информации (ресурсов автоматизированной информационной системы), при котором субъекты, имеющие соответствующие права доступа, могут беспрепятственно реализовывать их. К правам доступа относятся: право на чтение, изменение, копирование, уничтожение информации; права на изменение, использование и уничтожение ресурсов.
- *загрузочный вирус* – компьютерный вирус, записывающийся в первый сектор гибкого или жёсткого диска и выполняющийся при загрузке компьютера.
- *кейлогер* – программное обеспечение или аппаратное устройство, регистрирующее различные действия пользователя.
- *компьютерная атака* – целенаправленное несанкционированное воздействие на информацию, на ресурс автоматизированной информационной системы или получение несанкционированного доступа к ним с применением программных или программно-аппаратных средств.
- *компьютерный вирус* – это вредоносная программа, способная создавать вредоносные программы и (или) свои копии.

- *конфиденциальность информации* - необходимость предотвращения утечки (разглашения) какой-либо информации.
- *макровирус* – разновидность компьютерных вирусов, разработанных на макроязыках, встроенных в прикладные пакеты программного обеспечения.
- *полиморфный компьютерный вирус* – специальная техника, используемая авторами вредоносного программного обеспечения для снижения уровня детектирования вредоносной программы классическими антивирусными продуктами (не имеет сигнатур).
- *программное воздействие* – несанкционированное воздействие на ресурсы автоматизированной информационной системы, осуществляемое с использованием вредоносных программ.
- *руткит* – набор программных средств для обеспечения маскировки объектов, контроля событий, происходящих в системе, сбора данных о параметрах системы.
- *сетевая атака* – компьютерная атака с использованием протоколов межсетевого взаимодействия.
- *стелс-вирус* – вирус, полностью или частично скрывающий свое присутствие в системе, путем перехвата обращений к операционной системе, осуществляющих чтение, запись, чтение дополнительной информации о зараженных объектах.
- *файловый вирус* – компьютерный вирус, распространяющийся путем внедрения своего кода в тело исполняемых файлов.
- *целостность информации* – состояние информации (ресурсов автоматизированной информационной системы), при котором её (их) изменение осуществляется только преднамеренно субъектами, имеющими на него право.

1.2. Понятие защиты информации

В настоящее время существуют разные определения понятия «информация». Данный термин ассоциируется с понятиями: сведения, данные, знания, известие, сообщение и тому подобный. Особенность информации состоит в том, что проявляется она только при взаимодействии объектов, которые представляют собой организованную структуру. Обычно предполагается наличие двух объектов хранения информации – источника информации и приемника (потребителя) и нестационарного, то есть изменяющегося во времени, процесса ее передачи от первого ко второму посредством передающего и приемного устройств (рис. 1.1).

Источник информации – это субъект или объект, порождающий информацию и представляющий ее в виде *сообщения*. *Приёмник информации* – это субъект или объект, принимающий *сообщение* и способный его интерпретировать. В этих определениях под словосочетание «субъект или объект» означает, что источник и приемник информации могут быть как неодушевленными (бумага, технические устройства), так и одушевленными (человек, общество).

Процесс передачи информации связан с физическим процессом, несущим информацию о событии или состоянии объекта наблюдения и называемым *сигналом*. Однако одиночный сигнал не может содержать большой объем информации. Поэтому для передачи обычно используется последовательность сигналов, называемая *сообщением*. Следовательно, сообщение служит переносчиком информации.

Под *компьютерной безопасностью* понимают состояние защищенности вычислительных устройств и компьютерных сетей.

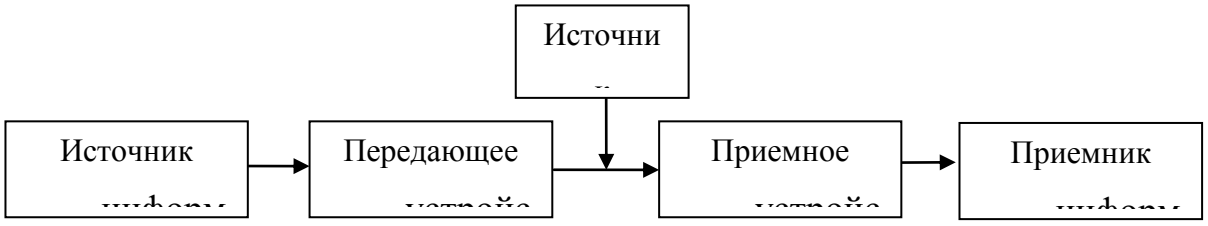

Рис. 1.1. Общая схема передачи информации.

Политика безопасности – это набор документированных норм, правил и практических приемов, регулирующих управление, защиту и распределение информации ограниченного доступа.

Угроза защиты информации – событие или действие, которое может вызвать изменение функционирования компьютерной системы, связанное с нарушением защищенности обрабатываемой в ней информации.

Атака – реализация угрозы защиты информации, которая заключается в поиске и использовании той или иной уязвимости.

Угрозы защиты информации могут быть разделены на естественные угрозы, не зависящие от деятельности человека, и искусственные угрозы, вызванные человеческой деятельностью.

Искусственные угрозы делятся на *непреднамеренные* (случайные) и *преднамеренные* (умышленные).

К непреднамеренным угрозам относятся ошибки в проектировании систем, в разработке программных средств, случайные сбои в работе аппаратных средств, ошибки пользователей, воздействие электромагнитных полей других устройств и так далее.

К преднамеренным угрозам относятся кроме несанкционированного доступа к ресурсам также и несанкционированные действия обслуживающего персонала, в том числе ослабление политики безопасности, несанкционированный доступ к ресурсам КС.

Преднамеренные угрозы защиты информации делятся:

- на угрозы нарушения конфиденциальности (утечки информации ограниченного доступа, в том числе параметров подсистемы защиты информации);
- угрозы нарушения целостности информации, хранящейся в компьютерной системе или передаваемой между ними;
- угрозы нарушения доступности информации, то есть отказа в обслуживании.

Утечка информации может происходить по косвенным каналам:

- использование подслушивающих устройств;
- дистанционное видеонаблюдение;
- перехват побочных электромагнитных излучений и наводок (ПЭМИН, TEMPEST, Transient Electromagnetic Pulse Emanation Standard)

и непосредственным каналам:

- хищение носителей информации;
- преднамеренное копирование файлов других пользователей;
- копирование носителей информации;
- сбор производственных отходов с информацией (бумажных и магнитных носителей);
- чтение остаточной информации после выполнения заданий других пользователей (областей оперативной памяти, удаленных файлов, ошибочно сохраненных временных файлов);
- намеренное использование для несанкционированного доступа к информации незаблокированных терминалов пользователей;
- маскировка под других пользователей путем похищения их идентифицирующей информации;
- незаконное подключение специальной регистрирующей аппаратуры к устройствам или линиям связи;
- злоумышленное изменение программ для выполнения ими несанкционированного копирования информации при ее обработке;
- злоумышленный вывод из строя средств защиты информации.

Наличие в системе значительного числа возможных каналов утечки информации обусловливает её уязвимость с точки зрения защиты информации.

Для обеспечения защиты информации необходимо применять системно-концептуальный подход, включающий целевую системность (защищенность информации рассматривается как составная неотъемлемая часть ее качества), пространственную системность (взаимосвязанность защиты информации во

всех элементах системы); временную системность (непрерывность защиты информации), организационную системность (единство организации всех работ по защите информации и управления компьютерными системами).

Компьютерную безопасность необходимо комплексно обеспечивать на всех этапах жизненного цикла системы с применением всех доступных методов и средств.

При передаче информации приемник информации должен не только принимать сообщение, но и его интерпретировать. Правилом *интерпретации сообщения* называется соответствие между сообщением и содержащейся в нем информацией. Данное соответствие может быть однозначным или неоднозначным. Неоднозначность интерпретации может быть вызвана передачей информации посредством различных сообщений для одного и того же приемника (например, передача письмом или шифротекстом) или передачей одного и того же сообщения для различных приемников (например, для зарегистрированного пользователя и нарушителя). Также в процессе передачи сообщения от источника информации к приемнику информация может искажаться. Это возможно в случае наличия помех (шумов) самих технических устройств (источника или приемника).

Существующие методы и средства обеспечения защиты информации делятся на методы и средства организационно-правовой защиты, инженерно-технической защиты, криптографические и программно-аппаратные.

1.3. Количество информации

Для измерения количества информации в теории кодирования принят энтропийный подход. Он основан на том, что получении информации связано с уменьшением разнообразия или неопределенности (*энтропии*) системы. Неопределенность понимается в смысле того, насколько мало известно наблюдателю о рассматриваемой системе. При получении информации энтропия (неопределенность) уменьшается, то есть система становится более упорядоченной.

Так как информация – содержание сообщения, в результате которого уменьшается энтропия (неопределенность) системы, то для того, чтобы измерить количество информации *I*, необходимо уметь вычислять ее энтропию *H*.

$$I = H_1 - H_2,$$

где H_1 - энтропия системы до получения информации;

H_2 - энтропия системы после получения информации.

Энтропию можно рассматривать и как количество информации, которое необходимо получить, чтобы система перестала быть неопределенной $H_2 = 0.$

Рассмотрим дискретную систему, имеющую конечное множество возможных состояний $\{s_i\}$, $i = \overline{1,n}$. Будем полагать, что все состояния системы различны. Множество состояний системы $S = \{s_1, s_2, \ldots, s_n\}$ называется ее *алфавитом*. А сами состояния s_i, $i = \overline{1,n}$ называются буквами, символами или знаками алфавита.

Рассматриваемая система может в каждый момент времени принимать одно из состояний s_i. Различные состояния могут возникать с различной *вероятностью* p_i. Однако для каждого состояния s_i вероятность p_i фиксирована. Так как система обязательно находится в каком-то из своих состояний, то сумма вероятностей возникновения какого-то состояния равна единице:

$$\sum_{i=1}^{n} p_i = 1.$$

Рассмотрим систему с равновероятными состояниями ($\forall i, j \;\; p_i = p_j = \frac{1}{n}$, $i, j = \overline{1,n}$). Чем в большем количестве возможных состояний может находиться система, тем меньше информации несет каждое состояние (больше оставшаяся энтропия системы). Если количество возможных состояний системы равно 1, то неопределенность системы отсутствует.

С другой стороны, если рассмотреть две независимые системы α и β с количеством равновероятных состояний n_α и n_β, то общая неопределенность двух систем (αβ) больше неопределенности каждой отдельно взятой и равна сумме их неопределенностей.

Таким образом, функция *H*, являющаяся мерой неопределенности системы, должна удовлетворять следующим условиям:

- она должна монотонно (непрерывно) возрастать с увеличением возможных состояний системы *n*: $H \in C^0$ (требование непрерывности), $\lim\limits_{n\to\infty} H(n) = \infty$;
- при $n = 1$ функция равна 0: $H(1) = 0$;
- должно выполняться требование аддитивности: $H(n_\alpha \cdot n_\beta) = H(n_\alpha) + H(n_\beta)$;

Так как количество состояний системы *S* положительно ($n \geq 1$), то указанным требованиям удовлетворяет логарифмическая функция с любым основанием, превышающим 1:

$$H = \log_k n.$$

Величина основания *k* определяет только масштаб или единицу измерения системы.

Указанная мера неопределенности – логарифмическая мера информации $k\log_2 n$ (мера Хартли) для систем с равновероятными состояниями была предложена американским ученым Ральфом Винтоном Лайоном Хартли в 1928 году. В зависимости от основания логарифма применяются следующие единицы измерения неопределенности:

- бит – $H = \log_2 n$;
- нат – $H = \ln n$;
- дит – $H = \lg n$.

Для системы с состояниями, возникающими с разной вероятностью, при вычислении энтропии следует также учитывать вероятность произошедшего

события. Действительно, в ситуации, когда система принимает менее вероятное состояние (с меньшим значением p_i), то информация о системе становится больше, чем при более вероятном состоянии. Действительно, если температура человека все время 36,6 градусов, то каждый новый момент времени несет небольшую информацию о свойствах организма. Когда же температура становится равной 38,3 градусам, то информация значительно возрастает (энтропия уменьшается).

Американский ученый Клод Шеннон обобщил понятие меры неопределенности H на случай системы S, когда состояния (символы алфавита) s_i имеют разную вероятность p_i:

$$H = -\sum_{i=1}^{n} p_i \log_k p_i .$$

Эта величина, характеризующая неопределенность, приходящуюся в среднем на одно состояние системы, называется *энтропией* дискретного источника информации.

Пример. Пусть система может находиться в одном из трех состояний, причем вероятности нахождения в первом и втором состояниях равны соответственно $p_1 = 0{,}4;\ p_2 = 0{,}1$. Найти энтропию системы.

Так как система может находиться в одном из трех состояний, а сумма вероятностей равна 1 ($p_1 + p_2 + p_3 = 1$), то вероятность нахождения системы в третьем состоянии равна $p_3 = 1 - 0{,}4 - 0{,}1 = 0{,}5$. Тогда $\log_2(p_1) = \log_2 0{,}4 = -1{,}32,\ \log_2(p_2) = \log_2 0{,}1 = -3{,}32,\ \log_2(p_3) = \log_2 0{,}5 = -1$ и энтропия H равна:

$$H = -(p_1 \cdot \log_2(p_1) + p_2 \cdot \log_2(p_2) + p_3 \cdot \log_2(p_3)) =$$
$$0{,}4 \cdot 1{,}32 + 0{,}1 \cdot 3{,}32 + 0{,}5 \cdot 1 = 1{,}36 \text{ бит.}$$

Ответ: энтропия системы равна 1,36 бит.

1.4. Представление информации в технических устройствах

Для обеспечения защиты информации необходимо рассмотреть формы представления данных в технических устройствах. Не нарушая общности,

можно рассматривать представление информации в дискретном виде, так как именно таким образом представленная информация обрабатывается компьютером и передается по различным линиям связи. Так как сообщение представляет собой последовательность сигналов (знаков некоторого алфавита), то при передаче данных возникает проблема распознавания знака. Требуется прочитать сообщение, то есть по полученным сигналам восстановить исходную последовательность знаков первичного алфавита. Для этого проводится анализ получаемой информации.

В общем случае информация, которая содержится в сообщении, может зависеть от того, в какой момент времени оно получено. Однако в ряде сообщений информация не зависит от конкретного времени его получения. Например, при передаче данных посредством вычислительной техники, с точки зрения принимающего устройства определенный знак всегда остается тем же знаком. То есть такая ситуация реализуется, когда вероятность встретить какой-либо знак в сообщении одинакова во все моменты времени. Обычно в этом случае вероятность равна относительной частоте этого знака во всей последовательности знаков. В таблицах приведены относительные вероятности употребления букв русского и английского языков.

Буква	Вероят-ность	Буква	Вероят-ность	Буква	Вероят-ность	Буква	Вероят-ность
Пробел	0,175	Р	0,04	Я	0,018	Х	0,009
О	0,09	В	0,038	Ы	0,016	Ж	0,007
Е, Ё	0,072	Л	0,035	З	0,016	Ю	0,006
А	0,062	К	0,028	Ь, Ъ	0,014	Ш	0,006
И	0,062	М	0,026	Б	0,014	Ц	0,004
Т	0,053	Д	0,025	Г	0,013	Щ	0,003
Н	0,053	П	0,023	Ч	0,012	Э	0,003
С	0,045	У	0,021	Й	0,01	Ф	0,002

Буква	Вероят-ность	Буква	Вероят-ность	Буква	Вероят-ность
Пробел	0.2	H	0.047	W	0.012
E	0.105	D	0.035	G	0.011
T	0.072	L	0.028	B	0.01
O	0.065	C	0.023	V	0.008
F	0.063	F	0.023	K	0.003
N	0.058	U	0.023	X	0.001
I	0.055	M	0.021	J	0.001
R	0.052	P	0.018	Q	0.001
S	0.052	Y	0.012	Z	0.001

Энтропия, приходящаяся в среднем на каждый знак русской буквы, составляет, согласно формуле (2.3), 4,36 бит, а на каждый знак английской буквы – 4,04 бита, французской – 3,96 бит, немецкой – 4,10 бит. Несовпадение значений энтропии для различных языков связано как с различным количеством букв языка, так и с различной вероятностью появления одних и тех же букв.

Сообщения, в которых вероятность появления каждого отдельного знака не меняется со временем, называются шенноновскими, а порождающий их отправитель – шенноновским источником.

Если сообщение является шенноновским, то набор знаков (алфавит) и информация, связанная с каждым из знаков, известны заранее. В этом случае интерпретация сообщения сводится к распознаванию конкретных знаков. Теория информации строится только для шенноновских сообщений.

В приведённых относительных вероятностях со средней информацией, приходящейся на буквы различных языков, были учтены вероятности появления букв в сообщениях. Однако если рассматривать не отдельные буквы, а их сочетания, то можно заметить, что некоторые из них вообще не встречаются. Например, в русском языке нет слов, содержащих пары *щц* и *фъ*. С другой стороны, некоторые сочетания встречаются более часто, а после, например, пары *пр* всегда следует гласная буква. Последовательность энтропий при учете возрастающего количества сочетаний является убывающей и стремится к некоторой величине $H_{\min}$, характеризующей минимальную неопределенность информации. Максимального значения энтропия достигает, как следует из ее свойств, при равной вероятности появления знаков алфавита.

Шеннон ввел величину, характеризующую рациональность применения символов алфавита, которую назвал избыточностью языка системы:

$$R = 1 - \frac{H_{\min}}{H_{\max}}.$$

Избыточность равна нулю только в случае независимости и равной вероятности знаков языка и максимальна при минимальном значении энтропии системы. Так как реальные системы могут находиться более чем в одном состоянии, то их энтропия не может быть равна нулю, и, следовательно, избыточность всегда принимает значения меньше единицы.

При измерении информации в битах, выражение для избыточности можно записать в следующем виде:

$$R = 1 - \frac{H_{\min}}{\log_2 n},$$

где n – количество знаков в алфавите языка.

Пример. Определить избыточность языка (в процентах), состоящего из четырех символов: *а*, *б*, *в*, *г*, если вероятности их появления составляют: $p_a = 0{,}6;\ p_a = 0{,}25;\ p_a = 0{,}1;\ p_a = 0{,}05.$

Энтропия системы равна:

$$H = -0{,}6 \cdot \log_2 0{,}6 - 0{,}25 \cdot \log_2 0{,}25 - 0{,}1 \cdot \log_2 0{,}1 - 0{,}05 \cdot \log_2 0{,}05 =$$
$$0{,}44 + 0{,}50 + 0{,}33 + 0{,}22 = 1{,}49 \text{ бит.}$$

$$R = 1 - \frac{1.49}{\log_2 4} = 1 - 0{,}745 = 0{,}255.$$

В процентах избыточность языка составляет $R = 0{,}255 \cdot 100\% = 25{,}5\%.$

Ответ: избыточность языка составляет 25,5%.

Избыточность языка показывает, какую долю лишней информации содержат тексты данного языка. Исследования Шеннона показали, что для английского языка $H_{\min} = 1{,}45$ бит. То есть его избыточность составляет $R = 1 - \frac{1{,}45}{4{,}75} = 0{,}69$, то есть 69%. Это означает, что английский текст можно сократить практически в три раза без ущерба для его содержательной стороны и выразительности. Однако это привело бы к значительному уменьшению разборчивости языка и ухудшению его распознавания при наличии шумов. При этом передача информации, связанная с битовыми

сообщениями, подразумевает равновероятное распределение информации, о чём будет сказано в разделе 3.

Чем больше избыточность, тем меньше требуется ресурсов линий связи при передаче сообщений, однако тем сложнее восстановить текст в случае ошибок при передаче. В этом смысле избыточность является определенной страховкой и гарантией разборчивости сообщений.

1.5. Задачи для самостоятельного решения

Задание. Имеются две системы, каждая из которых характеризуется двумя состояниями. Первая система находится в состоянии 1 с вероятностью 0,25, а вторая – 0,5. Вычислить энтропию систем? (0,81 бит, 1,0 бит)

Задание. Имеются два ящика (системы), в каждом из которых находится по 8 шаров двух цветов. В первом ящике 2 зеленых шара и 6 желтых шаров, а во втором – по 4 шара каждого цвета. Из каждого ящика вытаскивают по одному шару. Что можно сказать о неопределенностях опытов? (Так как энтропия (неопределенность) второй системы больше, чем первой, то предсказать исход опыта для второго ящика сложнее.)

Контрольные вопросы

1. Какие федеральные законы регламентируют деятельность по обеспечению защиты информации?
2. Что понимают под термином «информация»?
3. Что такое источник информации, приёмник информации?
4. Дайте определение защиты информации.
5. Чем определяется политика безопасности?
6. Что такое угроза защиты информации?
7. Приведите пример естественных и искусственных угроз защиты информации.
8. К какому типу угроз относятся непреднамеренные?
9. Перечислите виды преднамеренных угроз защиты информации.

10. Приведите пример непосредственных и косвенных каналов утечки информации.

11. Что такое энтропия?