

Теория строения скрытых систем

Коваленко Геннадий Александрович

1. Введение

Элементы сокрытия информации всегда неразрывно связаны со способами их передачи, даже само хранение информации есть частный случай её транспортировки [1, с.132]. Главной проблемой в создании соединения, со стороны классической криптографии, являлась проблема сдачи, передачи, сокрытия сеансового ключа, который в следствии служил для шифрования всей основной информации [2]. Главной же проблемой в создании соединения, со стороны современной криптографии, считается проблема доверия, искоренить которую представляется невозможным [1, с.267]. Связано это всё потому, что даже общаясь с получателем, ему необходимо доверять, чтобы он не разгласил в последствии полученную информацию, а следовательно и не нарушил её конфиденциальность. Тем не менее, всё это является крайней формой доверия, о которой следует заботиться только в последнюю очередь. Главной же проблемой здесь являются промежуточные субъекты (центры сертификации, серверы, узлы) хранения и транспортировки информации, от её создания до уничтожения, от отправления до её получения. В таких случаях, необходимо не только скрывать саму передаваемую информацию (объект транспортировки), но и её отправителя и получателя (субъектов транспортировки). На основе всего вышеперечисленного, строится система доверия, разрушить которую не представляется возможным, так как она (сама же система) защищает от других, более угрожающих, устрашающих, агрессивных, видов атак (например, MITM) и способствует более оптимизированному процессу выполнения задач, за счёт выделения клиентов и серверов в отдельные категории. Таким образом, построение новых систем, представляющих анонимность и безопасность пользователей, является очень кропотливым и расчётливым видом деятельности, учитывающим множество компромиссов, и основанным на факте уменьшения мощности доверия*. К системам подобного рода относятся анонимные сети и тайные каналы связи.

*Мощность доверия — количество узлов, участвующих в хранении или передаче информации, представляющую открытый характер для данных узлов. Иными словами, такие узлы способны читать, подменять и видоизменять информацию, так как для них она находится в открытом виде. Чем больше мощность доверия, тем выше предполагаемый шанс компрометации отдельных узлов, а следовательно, и хранимой на них информации. Принято считать одним из узлов получателя. Таким образом, нулевая мощность доверия будет возникать лишь в моменты отсутствия каких-либо связей и соединений. Если мощность доверия равна единице, это говорит о том, что связь защищена, иными словами, никто кроме отправителя и получателя информацией не владеют. Во всех других случаях мощность доверия будет больше единицы, что говорит о групповой связи (то-есть, о существовании нескольких получателей), либо о промежуточных узлах, способных читать информацию в открытом виде.

«Тогда лишь двое тайну соблюдают,

Когда один из них её не знает.»

(У. Шекспир, Ромео и Джульетта)

2. Анонимные сети

Скрытые, тайные, тёмные, анонимные сети – есть сети, грамотно соединяющие и объединяющие маршрутизацию вместе с шифрованием. Маршрутизация обеспечивает критерий анонимности, направленный на субъект, шифрование – критерии конфиденциальности, целостности, аутентификации, направленные на объект. Без маршрутизации легко определяются отправитель/получатель, без шифрования легко определяется передаваемое сообщение. Таким образом, только в совокупности этих двух свойств сеть может являться скрытой.

В современном мире большинство скрытых сетей представляют оверлейные соединения, иными словами соединения, которые основаны на уже существующей сети (например, сети Интернет). Но так или иначе, это не говорит, что скрытые сети не могут существовать сами по себе и быть однородной структурой, так как первоначальная архитектура может быть изначально нацелена на анонимность и безопасность. Именно по историческим причинам, современные сети имеют оверлейные уровни безопасности.

Любая анонимная сеть основывается либо на одноранговой, либо на гибридной архитектуре сети, исключая при этом многогранговую. Многогранговая сеть (или клиент-серверная архитектура) не может быть достаточно анонимной априори, потому как имеет открытую, прямую видимость и непосредственную реальность централизации. Централизация со стороны компьютерных сетей примитивна, централизация со стороны скрытых сетей неочевидна. В компьютерных сетях под централизацией понимается ограниченное количество серверов, способных обслуживать куда большее количество клиентов. Когда же говорим об анонимных сетях, то централизация в таких системах может иметь свойство приходящее и уходящее, тем самым, измеряя уровень централизации не в виде булевой логики, а в виде процентных соотношений от всей суммарности сети.

В одноранговых системах все пользователи однородны, имеют одинаковые возможности, могут представлять одни и те же услуги маршрутизации. Такие сети легко могут разворачиваться в локальных системах, но с осложнениями работают в глобальном пространстве на основе уже созданных соединений. Сами одноранговые сети могут быть разделены на две категории: децентрализованные и распределённые. Само отличие очень туманно со стороны терминологии, тем не менее, лучшим решением будет расслоение их свойств, для последующего повествования. Распределённые сети можно именовать “истинно децентрализованными” (хоть и противоречиво), где нельзя выделить какой-либо центр или узел связи сразу нескольких других узлов. Характерной чертой такой сети является тот факт, что каждый пользователь одновременно соединяется напрямую сразу с несколькими другими. Децентрализованные же сети можно именовать “слабо централизованными” (хоть также и противоречиво), где к одним узлам сети подключаются сразу несколько других узлов. Характерной чертой такой сети является тот факт, что появляются “неофициальные” узлы, часто используемые другими узлами (узлы-серверы) в качестве последующей маршрутизации.

Гибридная система объединяет свойства многогранговых и одноранговых архитектур, пытаясь взять и удержать как можно больше положительных и меньше отрицательных качеств. Плюсом многогранговых архитектур являются некоторые свойства централизации, как например возможность разделения логики на серверную и клиентскую, более быстрая и/или статичная скорость маршрутизации. Плюсом одноранговых архитектур являются некоторые свойства децентрализации, как собственно возможность достижения и построения анонимности. Гибридные сети не всегда можно однозначно определить, иногда они выглядят и/или ведут себя сродни одноранговым, иными словами используют минимальное количество свойств многогранговой архитектуры, или наоборот, могут иметь вид многогранговой архитектуры и даже

перерасти в неё полностью. Из всего этого многообразия, гибридные сети являются самой противоречивой системой.

Сама анонимность также не является чем-то однородным, точно определённым, абсолютным. Её можно рассматривать как некую градацию, набор стадий, поэтапность. Всего же существует шесть стадий и четыре вида анонимности (2,3,4,5 стадии).

0. Нулевая стадия представляет абсолютную анонимность. Примером является отсутствие соединений, а следовательно, и сети как таковой. $|A| = -\infty^*$.

1. Первая стадия характеризуется абсолютным отсутствием анонимности. Примером может являться прямое соединение между двумя пользователями. $|A| = 0$.

2. Вторая стадия игнорирует анонимность клиент-сервер и обеспечивает только анонимность клиент-клиент. Иными словами, сервер обладает достаточной информацией о клиентах (например, знание IP-адреса, личных данных, интересов). При этом сами клиенты не знают друг друга и потому являются анонимами. Типичным примером могут служить форумы, социальные сети, иначе говоря, большинство сайтов и приложений, построенных на основе клиент-серверной архитектуры. Из этого примера следует, что данная архитектура способна выдавать минимальную мощность анонимности, а именно $|A| = 1$. Второй этап является псевдо-анонимностью.

3. Третья стадия заменяет сетевой адрес криптографическим. Под заменой подразумевается неполное, частичное преобразование, так как сетевой адрес нельзя никак полностью скрыть, удалить, он будет существовать всегда и постоянно. Примером могут являться криптовалюты и блокчейн платформы, где известны лишь криптографические адреса клиентов (публичные ключи, хеши публичных ключей) [3]. Основным и главным отличием третьей стадии анонимности от второй, является тот факт, что сеть представляет собой гибридный, а следовательно, и раздробленный, неопределённый характер поведения узлов сети, имеющих разные интересы. Из этого образуется и повышенная мощность анонимности, так как каждый узел может перенаправлять, распространять, транслировать информацию другим узлам, не заботясь о сохранении сетевого адреса клиента. $0 < |A| \leq N$, где N - количество узлов в сети. Тем не менее, маршрутизация не играет здесь роли анонимизации субъекта, а необходима лишь для распространения объекта. Получателя в такой системе как такового не существует и потому все действия ориентированы на сохранение информации на всех (или большинстве) узлах сети. Поэтому, третью стадию можно охарактеризовать игнорированием анонимности со стороны субъекта и её сохранением в передаваемом объекте. На основе этой стадии также могут работать тайные каналы связи.

4. Четвёртая стадия изменяет способ маршрутизации, ограничивая доступ узлов не только к объекту (передаваемой информации), но и к субъектам (отправителю и получателю). Мощность анонимности такого этапа чаще всего статична, тем самым $\lim_{|A| \rightarrow C}$, где C - заданная по умолчанию константа. Примером данного этапа являются сами скрытые сети, подобия Tor (луковая маршрутизация), I2P (чесночная маршрутизация), RetroShare (turtle маршрутизация) и т.д.

5. Пятая стадия повышает мощность анонимности до теоретического максимума за счёт "слепой" маршрутизации (заливки) [4, с.398]. $\lim_{|A| \rightarrow N}$, где N - количество узлов в сети. В

таком способе не указывается получатель, а криптографический адрес отправителя зашифровывается на моменте отправления. Благодаря этому свойству, безопасность приложений этого рода полностью (или в большей мере) зависит от клиентской части. Пятый этап является также представителем скрытых сетей, в качестве примера которых можно привести BitMessage.

Стоит также заметить, что шифрование, определяемое анонимностью, появляется частично лишь в моменты третьей стадии, в то время как во второй - само шифрование является добавочным, дополнительным и второстепенным свойством, служащим лишь и только для защиты клиент-серверной коммуникации.

В третьей стадии, под шифрованием чаще всего понимается только подписание информации и генерация асимметричных ключей. Таким образом, как такового шифрования на этой стадии ещё не происходит, но рождается очень важная концепция - возможность идентификации в одноранговых и гибридных системах, а следовательно, целостность и аутентификация самой передаваемой информации [5, с.223].

На четвёртом этапе добавляется шифрование самой информации (иными словами рождается свойство конфиденциальности), так как теперь, в отличие от третьего этапа, информация представляет собой суть секретного, скрытого, тайного, а не открытого и общего объекта. Благодаря шифрованию, появляется возможность грамотно использовать маршрутизацию для повышения уровня анонимности в системах с получателем.

В четвёртой стадии существуют так называемые входные (entry) и выходные (exit) узлы, которые имеют дополнительную информацию либо об отправителе, либо о получателе. Так, например, входные узлы имеют IP-адрес отправителя, выходные - IP-адрес получателя. Всё это есть секретная информация, которая так или иначе может навредить анонимности. Решений данной проблемы несколько.

- Первый способ нацелен на увеличение мощности анонимности минимум до трёх. Это обуславливается тем фактом, что даже если входные и выходные узлы окажутся скомпрометированными и будут обладать общей целью, то есть вычислением отправителя и получателя, то им помешает промежуточный узел. Точно также, если входной и промежуточный узел будут скомпрометированы, то им помешает выходной узел для полной деанонимизации субъектов. Всё это выходит из того правила, что мощность анонимности в таких случаях не падает ниже двух. Таким образом, одновременное выявление и отправителя, и получателя является сложной задачей.
- Второй же способ нацелен на переход к пятой стадии - крайней форме анонимности. В таком случае, целью является сокрытие, удаление, исключение всех возможных связей между отправляемой информацией (объектом) и самим отправителем/получателем (субъектом). На таком этапе, существование входных и выходных узлов возможно лишь при условии полного контроля сети, по причине сложного обнаружения и последующего построения связей между объектом и субъектом.

Основным и главным отрицательным свойством пятой стадии анонимности является линейное увеличение нагрузки на сеть $O(N)$ со стороны всех пользователей в ней участвующих. Так, например, если сеть состоит из N узлов, то каждый узел должен будет обрабатывать $N-1$ запросов от других узлов. Время жизни пакета (TTL) на пятой стадии не является решением данной проблемы, по причине появления новых связей между отправителем и передаваемой информацией, что, следовательно, ведёт к переходу на четвёртый этап.

Большинство атак, направленных на скрытые сети, представляют способы деанонимизации субъектов (как наиболее лёгкий способ), нежели попытки раскрытия, взлома, дешифрования объектов (как наиболее сложный). Так например, достаточно сильной и сложно искоренимой атакой на одноранговые (а следовательно, и на гибридные) сети является атака Сивиллы. Она базируется на том факте, что главным способом анонимности является элемент маршрутизации, который обеспечивается за счёт передачи информации посредством нескольких узлов. С одной стороны, сутью атаки является замена несвязанных между собой узлов, на узлы подчинённые одному лицу, либо группе лиц с общими интересами, тем самым, атака ориентируется на $(\lim_{|A| \rightarrow 1})$ уменьшение мощности анонимности до единицы. С другой стороны, в некоторых видах сетей, с мощностью доверия больше единицы, атака может вредить и целостности передаваемой информации, иными словами подменять и видоизменять её. При повышении количества узлов несвязанных между собой в сети, повышается и сложность реализации атаки Сивиллы, за счёт более равномерного распределения узлов. Из этого также следует, что мощность анонимности будет стремиться к своему константному значению $\lim_{|A| \rightarrow C}$ (если анонимность представляет четвертую стадию) или к количеству несвязанных узлов $\lim_{|A| \rightarrow N}$ (если анонимность представляет пятую стадию). Всё это связано с тем, что атакующие узлы будут конкурировать с обычными узлами за возможность быть посредниками между субъектами передаваемой информации. Чем больше несвязанных узлов и лучше алгоритм распределения, тем меньше вероятность осуществления данной атаки.

* Мощность анонимности — количество узлов выстроенных в цепочку и участвующих в маршрутизации информации, при этом не будучи никак связанными между собой общими целями и интересами. Из этого следует, что многоранговая архитектура по умолчанию имеет мощность анонимности равной единице (вне зависимости от количества серверов). Нулевая мощность анонимности возникает лишь в моменты прямого соединения. Если соединения не существует, то мощность анонимности неопределенна.

Мощность анонимности:

$$|A| = \begin{cases} 0, & |S| = 0 \\ 1, & K = 0, \\ K, & K \neq 0 \end{cases}$$

где $S = S_C(N)$,

N - множество узлов, расположенных в сети,

S_C - функция выборки множества узлов, участвующих в маршрутизации,

$K = |S| - |E(S)|$,

E - функция выборки множества узлов, подчиняющихся одному лицу или группе лиц с общими интересами.

3. Тайные каналы связи

Скрытые, секретные, невидимые, тайные каналы связи - есть соединения, располагаемые в заведомо замкнутом, незащищённом, враждебном окружении и имеющие характеристики безопасной передачи информации. Анонимность, подобная анонимности в скрытых сетях, в тайных каналах связи отсутствует, так как пользователь уже находится в сети, следовательно, и является авторизованным её участником. Отличительной и своеобразной особенностью анонимности в тайных каналах, является не сокрытие субъектов информации, как это было в

скрытых сетях, а сокрытие самой информации (что отсылает нас в той или иной мере к третьей стадии анонимности). Скрытые каналы связи могут быть основаны на стеганографических [6] и/или криптографических свойствах. Их объединение порождает более безопасные и защищённые секретные каналы связи.

Тайные каналы связи, использующие стеганографию, всегда имеют некий контейнер, в который помещается истинное сообщение. Под контейнером может пониматься ложное, неявное, сбивающее с пути сообщение, которое чаще всего носит нейтральный характер. Из этого также следует, что в зависимости от размера контейнера, зависит и размер самого исходного сообщения, тем самым, стеганографический подход рассчитан на сообщения малых размеров и мало пригоден для передачи целых файлов. Примером контейнера может служить изображение, аудио-запись, видео-файл, то есть всё, что может хранить дополнительную или избыточную информацию, которая останется незаметной для человеческих глаз и ушей. Одним из примеров сокрытия информации может служить замена каждого старшего бита в изображении, битом исходного сообщения. Таким образом, если размер изображения (то есть, контейнера) будет равен 2MiB, то максимальный размер исходного сообщения не будет превышать 256KiB.

Тайные каналы связи, использующие криптографию, можно охарактеризовать третьей стадией анонимности. Иными словами, так как отсутствует структура маршрутизации, то 4-5 стадии являются недоступными. Если тайный канал разворачивается в заведомо замкнутой и незащищённой, но всё же сети, то это говорит о том, что стадия анонимности не меньше второй. Сами же скрытые каналы данного вида используют идентификацию по криптографическим адресам, а не адресам, заданными системой по умолчанию (никнеймом, телефоном и т.д.), следовательно, стадия анонимности тайных каналов определяется третьим этапом. Таким образом, секретные каналы улучшают безопасность уже выстроенной системы в неизменном для неё виде, используя лишь и только её базис.

Есть ещё один, и достаточно интересный, способ сокрытия информации, подобен стеганографическим, но только с использованием криптографии [7, с.720]. В таком методе истинная информация скрывается в цифровой подписи ложного сообщения.

Использование стеганографии, вместе с криптографией, может помочь в случаях, когда имеется вероятность или возможность нахождения скрытого сообщения в контейнере. Тем самым, даже если исходное сообщение было найдено, оно будет иметь зашифрованный вид. Здесь стоит учитывать тот факт, что при шифровании размер информации увеличивается (добавляется хеш, подпись, текст дополняется до блока), а из этого уже следует, что максимальный размер исходного сообщения уменьшается.

Теоретически, тайные каналы связи также могут находиться и в других скрытых каналах, либо анонимных сетях (секретные каналы могут быть воссозданы совершенно в разных системах и ситуациях), тем не менее, подобный подход является очень сомнительным (по причине избыточности накопленных слоёв шифрования), специфичным (по причине редкости практического использования) и затрачиваемым (по причине уменьшения производительности программ, уменьшения ёмкости контейнеров).

4. Протокол безопасной передачи информации

Из всего вышесказанного можно создать легковесный, примитивный, но при этом и безопасный протокол передачи информации, основанный на пятой стадии анонимности, потому как обеспечивает полное сокрытие субъектов в передаваемом объекте при помощи методов шифрования. Протокол является самодостаточным, цельным, монолитным, может быть применим к скрытым сетям и тайным каналам связи [1, с.58][7, с.80].

Участники протокола:

А - отправитель,

В - получатель.

Шаги участника А:

1. $K = G(N), R = G(M)$,

где G - функция-генератор случайных байт,
 N, M - количество байт для генерации,
 K - сеансовый ключ шифрования,
 R - случайные набор байт.

2. $H_M = H(R, M, PubK_A, PubK_B)$,

где H_M - хеш сообщения,
 H - функция хеширования,
 M - исходное сообщение,
 $PubK_A, PubK_B$ - публичные ключи отправителя и получателя.

3. $C_R = [E(PubK_B, K), E(K, PubK_A), E(K, M), E(K, R), H_M, E(K, S(PrivK_A, H_M)), W(C, H_M)]$,

где C_R - зашифрованное сообщение,
 E - функция шифрования,
 S - функция подписания,
 W - функция подтверждения работы,
 C - сложность работы,
 $PrivK_A$ - приватный ключ отправителя.

Шаги участника В:

4. $W(C, H_M) = P(C, W(C, H_M))$,

где P - функция проверки работы.
Если \neq , то протокол прерывается.

5. $K = D(PrivK_B, E(PubK_B, K))$,

где D - функция расшифрования,
 $PrivK_B$ - приватный ключ получателя.

Если расшифрование неверно, то протокол прерывается.

6. $PubK_A = D(K, E(K, PubK_A))$.

7. $H_M = V(PubK_A, D(K, S(PrivK_A, H_M)))$,

где V - функция проверки подписи.
Если \neq , то протокол прерывается.

8. $H_M = H(D(K, E(K, R)), D(K, E(K, M)), PubK_A, PubK_B)$,

Если \neq , то протокол прерывается.

Данный протокол игнорирует способ получения публичного ключа от точки назначения. Это необходимо по причине того, чтобы протокол был встраиваемым и мог внедряться во множество систем, включая одноранговые сети, не имеющие центров сертификации, и тайные каналы связи, имеющие уже установленную сеть по умолчанию.

Безопасность протокола определяется в большей мере безопасностью асимметричной функции шифрования, так как все действия сводятся к расшифрованию сеансового ключа приватным ключом. Если приватный ключ не может расшифровать сеансовый, то это говорит о том факте, что само сообщение было зашифровано другим публичным ключом и потому получатель также есть другой субъект. Функция хеширования необходима для проверки целостности отправленных данных. Функция проверки подписи необходима для аутентификации отправителя. Функция проверки доказательства работы необходима для предотвращения спама.

Шифрование подписи сеансовым ключом является необходимым, так как взломщик протокола, для определения отправителя (а именно его публичного ключа) может составить список уже известных ему публичных ключей и проверять каждый на правильность подписи. Если проверка приводит к безошибочному результату, то это говорит об обнаружении отправителя.

Протокол пригоден для многих задач, включая передачу сообщений, запросов, файлов, но не пригоден для передачи поточной информации, подобия аудио звонков и видео трансляций, из-за необходимости подписывать и подтверждать работу, на что уходит много времени. Иными словами, протокол работает с конечным количеством данных, размер которых заведомо известен и обработка которых (то есть, их использование) начинается с момента завершения полной проверки.

Код шифрования на языке Go выглядит так:

```
import (
    "bytes"
    "encoding/hex"
)
func Encrypt(sender *PrivateKey, receiver *PublicKey, data []byte) *Package {
    var (
        session = GenerateBytes(N)
        rand     = GenerateBytes(M)
        pubsend  = PublicKeyToBytes(&sender.PublicKey)
        hash     = HashSum(bytes.Join(
            [][]byte{
                rand,
                data,
                pubsend,
                PublicKeyToBytes(receiver),
            },
            []byte{}),
        ))
        sign = Sign(sender, hash)
    )
    return &Package{
        Head: HeadPackage{
            Rand: hex.EncodeToString(EncryptS(session, rand)),
            Sender: hex.EncodeToString(EncryptS(session, pubsend)),
            Session: hex.EncodeToString(EncryptA(receiver, session)),
        },
        Body: BodyPackage{
            Data: hex.EncodeToString(EncryptS(session, data)),
            Hash: hex.EncodeToString(hash),
            Sign: hex.EncodeToString(EncryptS(session, sign)),
            Npow: ProofOfWork(hash, C),
        },
    }
}
```

Функции `GenerateBytes`, `HashSum`, `PublicKeyToBytes`, `Sign`, `EncryptS`, `EncryptA`, `ProofOfWork`, а также типы данных `PrivateKey`, `PublicKey`, `Package` являются кастомными, самописными, представляющими интерфейс, а следовательно и не реализованными по умолчанию в языке Go [8].

Для улучшения эффективности, допустим при передаче файла, программный код можно изменить так, чтобы снизить количество проверок работы в процессе передачи, но с

первоначальным доказательством работы на основе случайной строки (полученной от точки назначения), а потом и с накопленным хешем из n -блоков файла, для i -ой проверки. Таким образом, минимальный контроль работы будет осуществляться лишь $\lceil M/nN \rceil + 1$ раз, где M - размер файла, N - размер одного блока. Если доказательство не поступило или оно является неверным, то нужно считать, что файл был передан с ошибкой и тем самым запросить повреждённый или непроверенный блок заново.

5. Список литературы

1. Шнайер, Б., Фергюсон, Н. Т. Практическая криптография / Б. Шнайер, Н. Т. Фергюсон. - М.: Издательский дом «Вильямс», 2005. - 420 с.
2. Диффи. В., М. Хеллман. Новые направления в криптографии. Режим доступа: <https://ee.stanford.edu/~hellman/publications/24.pdf> (дата обращения: 19.12.2020).
3. Накамото, С. Биткойн: система цифровой пиринговой наличности [Электронный ресурс]. — Режим доступа: https://bitcoin.org/files/bitcoin-paper/bitcoin_ru.pdf (дата обращения: 19.12.2020).
4. Таненбаум, Э., Уэзеролл, Д. Компьютерные сети / Э. Таненбаум, Д. Уэзеролл. — СПб.: Питер, 2017. - 960 с.
5. Рябко, Б. Я., Фионов, А. Н. Криптография в информационном мире / Б. Я. Рябко, А. Н. Фионов. - М.: Горячая линия - Телеком, 2019. - 300 с.
6. Шелухин, О.И., Канаев, С.Д. Стеганография. Алгоритмы и программная реализация / О.И. Шелухин, С.Д. Канаев. — М.: Горячая линия - Телеком, 2018. - 592 с.
7. Шнайер, Б. Прикладная криптография. Протоколы, алгоритмы и исходные коды на языке С / Б. Шнайер. — СПб.: ООО «Альфа-книга», 2018. - 1040 с.
8. Донован, А.А., Керниган, Б.У. Язык программирования Go / А.А. Донован, Б.У. Керниган. — М.: ООО «И.Д. Вильямс», 2018. - 432 с.