Метод автоматизированной идентификации признаков протоколов сетей передачи данных

Method of automated identification of data transmission network protocol features

Дементьев / Dement'ev V.
Владислав Евгеньевич
(dem-vlad@rambler.ru)
доктор технических наук, доцент.
ФГКВОУ ВО «Военная академия связи имени Маршала Советского Союза С. М. Буденного» МО РФ (ВАС им. С. М. Буденного),
заместитель начальника кафедры.
г. Санкт-Петербург

Чулков / Chulkov A.
Александр Анатольевич
(sir.alexanderchulkov@yandex.ru)
ВАС им. С. М. Буденного,
адъюнкт.
г. Санкт-Петербург

Ключевые слова: идентификация – identification; признак – feature; протокол – protocol; уязвимость – vulnerability; оценка – assessment; защищенность – security; автоматизация – automation.

Рассматривается подход к разработке метода автоматизированной идентификации признаков протоколов сетей передачи данных, имеющего в своей основе подготовку спецификаций и массивов сигнатур для оценки защищенности протоколов. В основе метода лежит идея достоверной идентификации сигнатур признаков уязвимостей протоколов сети передачи данных, позволяющая оценить их защищенность от возможных воздействий, а также на предмет ошибок реализации, в том числе и за счет незадекларированных возможностей. Наборы тестовых сигнатур формируются на основе результатов анализа трафика статически (до начала работы) или динамически (в процессе реализации метода).

In the article, the authors consider an approach to the development of a method for automated identification of features of data transmission network protocols, which is based on the preparation of specifications and signature arrays for evaluating Protocol security. The method is based on the idea of reliable identification of signatures of signs of vulnerabilities in data transmission network protocols, which makes it possible to assess their security from possible impacts, as well as for implementation errors, including due to undeclared features. Sets of test signatures are generated based on the results of traffic analysis, either statically (before the start of work) or dynamically (during the implementation of the method).

Введение

В настоящее время в мире насчитывается несколько тысяч различных спецификаций протоколов обмена данными. Кроме того, крупные производители IT-оборудования разрабатывают свои собственные версии протоколов или обновляют уже существующие.

Примитивы протоколов, программные сущности, их реализующие, обладают огромным количеством индивидуальных особенностей, которые выделяют их в общем пространстве передаваемых единиц обмена данными. В данном случае речь идет не об алгоритмических или программных особенностях, а об сигнатурных и семантических признаках, несущих какую-то служебную нагрузку. Наличие этих особенностей говорит о возможности однозначной идентификации протоколов в процессе служебного или информационного обмена данными. В общем случае для осуществления этой процедуры необходимо выполнить анализ сигнатур признаков в формате бинарного кода, передаваемого в виде кадров или пакетов.

В нашем случае под анализом бинарного кода будем понимать проверку корректности функционирования протоколов и отсутствие погрешностей или аномалий соответствующих значений полей заголовков кадров или пакетов данных. Наличие подобных признаков [2–6] расценивается как потенциальные уязвимости, обусловленные либо особенностями реализации, либо наличием незадекларированных разработчиками возможностей, либо их промахами (уязвимостями), эксплуатируемыми злоумышленниками. Проявлением этих ошибок является аварийное (нештатное) функционирование протокола, приводящее к нарушению обмена данными или функционирования СПД в целом.

В современных технических решениях и подходах [1, 5, 7] анализ и оценка двоичных последовательностей выполняется по следующим направлениям:

– статический или динамический анализ бинарного кода;

– инверсионный анализ данных (reverse engineering);

– применение проверочных сигнатур (стрессовое тестирование).

Широко известен подход для оценки защищенности на основе применения технологии фазинга (fuzzing) [1]. Опыт внедрения данного подхода позволяет утверждать

о возможности его применения и для оценки протоколов СПД. Одними из его плюсов являются отсутствие необходимости обновления баз сигнатур воздействий на протоколы СПД и возможность получения результатов за короткий промежуток времени. Для целей проводимого исследования целесообразно рассматривать мутационный и генерационный подходы, лежащие в основе фазинга. В первом предполагается хаотичное изменение входных сигнатур, а во втором – создание входных сигнатур на основе определенных правил. В нашем случае под мутацией предлагается понимать модификацию (изменение, трансформацию, видоизменение) полей пакетов и их содержимого как в сигнатурной, так и в семантической плоскостях двоичного кода, содержащегося в значениях полей заголовка, данных или концевика пакета или кадра.

Таким образом, опираясь на принципы модификации идентифицируемых признаков, авторами разработан метод автоматизированной идентификации признаков в протоколах сетей передачи данных (АИПП СПД), обладающий следующими свойствами:

– может применяться в условиях отсутствия заранее подготовленных данных об идентификационных признаках воздействий на протоколы СПД, их уязвимостей или потенциальных угроз;

– предназначен для идентификации признаков, относящихся к категории «недекларируемых возможностей», которые, как правило, трудно выявить, основываясь на формальных методах анализа;

– относится к категории «динамических» (идентификация признаков осуществляется в процессе обмена данными, анализируется бинарная последовательность, передаваемая в виде трафика или функционирующего в СПД процесса);

– приемлем к применению относительно большого числа протоколов СПД (прикладных, сетевых, специализированных, технологических и т. п.);

- универсален для использующихся программно-аппаратных платформ и технологий построения СПД.

Разработанный метод предназначен для преобразования многопараметрических данных в модель идентификации сигнатур признаков.

Модель идентификации сигнатур признаков

Одним из элементов метода АИПП СПД является модель идентификации сигнатур признаков.

Условимся, что под оцениваемым протоколом (ОП) будем понимать протокол, который анализируется на предмет идентификации сигнатур признаков. Признаки ОП, как правило, имеют индивидуальные сигнатуры. В случае если идентифицируемый признак в протоколе находится в данных пакета (заголовок или поле данных) и имеет формализованную структуру, то его сигнатура состоит из последовательности бит. Тогда, с учетом [2–6], сигнатура признака – иерархически организованная взаимосвязанная последовательность данных, позволяющая идентифицировать их между собой на общем пространстве сигнатур. Однотипные сигнатуры обладают взаимосвязанными данными или признаками. Минимальная сигнатура будет иметь идентификатор размером один бит и относиться к категории – элементарная сигнатура. Тогда каждая сигнатура, за исключением элементарной, состоит из вложенных сигнатур, что обусловлено иерархией признаков ОП. Сигнатура каждого признака представляется собой последовательность (строку) бит. Длину каждой такой последовательности назовем размером сигнатуры [4].

Тип сигнатуры – свойство (атрибут) признака протокола, однозначно задающее множество значений, которые могут принимать сигнатуры данного типа.

Идентификатор сигнатуры признака протокола – совокупность инструкций и алгоритмов ОП, осуществляющего обработку полей пакета, содержащего сигнатуры признака протокола. Идентификатор однозначно классифицирует сигнатуру.

Предположим, что каждому признаку ОП соответствует некоторая сигнатура. Тогда:

Ω – множество сигнатур всех ОП;

Ω^* – множество сигнатур ОП из множества Ω;

$\Omega_n \subset \Omega$ – множество известных сигнатур всех ОП;

T – множество типов сигнатур всех ОП;

$byte \in T$ – тип элементарной сигнатуры;

$\Lambda \notin T$ – символ, обозначающий невозможность идентификации признака (ОП обрабатывает признаки, содержащие неизвестные сигнатуры или определение типов сигнатур затруднено);

$\Omega_t \subset \Omega$ – множество типовых сигнатур;

Θ – множество идентификаторов сигнатур признаков ОП;

$\Delta \notin \Theta$ – символ, означающий, что идентификатор сигнатуры признака неизвестен (отсутствует полное описание алгоритмов ОП);

$B = \{0,...,255\}$ – множество значений элементарной сигнатуры;

$B^n = \{(b_1,...,b_n)\}$ – множество всех длин сигнатур $n \geq 1$, где $b_i \in B$;

$B^* = \bigcup_{n \geq 1} B^n$ – множество всех сигнатур конечной длины из элементов в B.

Тогда функция $\lambda : \Omega \to N$ – задает размер сигнатур, где N – множество натуральных чисел, а функция $\nu : \Omega \to B^*$ – задает значение сигнатур. При этом каждое значение сигнатуры $\alpha \in \Omega$ задается последовательностью из множества $B^{\lambda(\alpha)}$.

Следовательно $\tau : \Omega \to T \bigcup \{\Lambda\}$ – тип сигнатуры и его функция;

$\omega : \Omega \to \Theta \bigcup \{\Delta\}$ – функция идентификатора каждой сигнатуры;

$\sigma : \Theta \to N$ – функция, задающая количество уровней иерархии для каждого идентификатора.

Тогда справедливы следующие условия:

– для двух сигнатур $\alpha_1, \alpha_2 \in \Omega$ существует равенство $\alpha_1 = \alpha_2$, если выполняется равенство $\omega(\alpha_1) = \omega(\alpha_2)$;

– сигнатура $\alpha \in \Omega$ известна ($\alpha \in \Omega_n$), если выполняется условие $\omega(\alpha) \neq \Delta$;

– если у сигнатуры $\alpha \in \Omega$ не определен тип ($\tau(\alpha) = \Lambda$), то она неизвестна ($\alpha \in \Omega_n$, $\omega(\alpha) = \Delta$);

– если у сигнатуры $\alpha \in \Omega$ определен тип ($\alpha \in \Omega_t$), то верно выражение $\tau(\alpha) \neq \Delta$;

– у известных сигнатур определен их тип, таким образом выполняется выражение $\Omega_n \subset \Omega_t$.

Установим на множестве типов сигнатур T отношение частичного порядка «$\leq$» (отношение «подтип сигнатуры»). При этом тип $t_1 \in T$ является подтипом типа $t_2 \in T(t_1 \leq t_2)$, если сигнатура типа t_2 содержит сигнатуру типа t_1. Если $t_1 \leq t_2$ и $t_1 \neq t_2$, тогда $t_1 < t_2$.

Множество базовых типов $T_0 \subset T$ является подмножеством множества типов T, такое, что выполняется условие $byte \in T$ и для любого $t \in T_0$ не существует типа $s \in T \setminus \{byte\}$ такого, что справедливо неравенство $s < t$. Тогда множество производных типов введем как $T_1 = T \setminus T_0$.

Примерами базовых типов могут быть «заголовок» или «поле».

Следовательно, если сигнатура $\alpha \in \Omega$ является сигнатурой базового типа, то справедливо выражение $\tau(\alpha) \in T_0$. Тогда сигнатура α является сигнатурой производного типа, если справедливо выражение $\tau(\alpha) \in T_1$.

Примем функцию иерархии типов как $H_T : T \to 2^T$, сопоставляющую каждому типу $t \in T\{byte\}$ множество типов $H_T(t) \subset T$ и удовлетворяющую условию: если $r \in H_T(t)$, то $r < t$ и нет $s \in T$ такого, что $r < s < t$. Тогда $H_T(byte) = \emptyset$.

Установим на множестве сигнатур Ω отношение частичного порядка «$\leq$» (отношение «элемент сигнатуры»). При этом сигнатура $\alpha \in \Omega$ представляет собой элемент сигнатуры $\beta \in \Omega(\alpha \leq \beta)$ если сигнатура β входит в состав сигнатуры β. Если $\alpha \leq \beta$ и $\alpha \neq \beta$, тогда $\alpha < \beta$ и соблюдается условие: если $\alpha \leq \beta$, то $\tau(\alpha) \leq \tau(\beta)$ и $\lambda(\alpha) \leq \lambda(\beta)$.

Предположим, что две различные сигнатуры $\alpha_1, \alpha_2 \in \Omega$, $\alpha_1 \neq \alpha_2$ имеют одинаковый тип ($\tau(\alpha) = \tau(\beta)$), но их идентификаторы различны $\omega(\alpha_1) \neq \omega(\alpha_2)$. Это справедливо в том случае, если сигнатура – значение поля α_1, входящая в состав сигнатуры «адресное поле пакета TCP», имеет идентификатор, отличный от идентификатора сигнатуры – значение поля α_2, входящей в состав сигнатуры «адресное поле пакета UDP». В данном случае, для этих сигнатур выполняются условия $\tau(\alpha_1) = \tau(\alpha_2)$ и $\omega(\alpha_1) \neq \omega(\alpha_2)$. Между тем каждая из рассматриваемых сигнатур относится к категории либо известной, либо неизвестной.

Представим $H_\Omega : \Omega \to \Omega^*$ – как функцию иерархии сигнатур, сопоставляющую каждой сигнатуре $\alpha \in \Omega$, где $\tau(\alpha) = byte$, конечную последовательность сигнатур $H_\Omega(\alpha) = (\alpha_1, ..., \alpha_m) \in \Omega^m, m \geq 1$, удовлетворяющих условиям: $\alpha_i < \alpha$ и не существует сигнатуры $\beta \in \Omega$, где $\alpha_i < \beta < \alpha (i = 1, ..., m)$. Если $\tau(\alpha) = byte$, то $H_\Omega(\alpha)$ – пустая последовательность (нулевой длины).

Если для сигнатуры $\alpha \in \Omega$ справедливо отношение $H_\Omega(\alpha) = (\alpha_1, ..., \alpha_m)$, тогда, на основе вышесказанного, выполняется равенство $H_T(\alpha) = \{\tau(\alpha_1), ..., \tau(\alpha_m)\}$.

Возможен вариант, когда типы и значения самой сигнатуры $\alpha \in \Omega$ и входящих в нее сигнатур не определены [4]. Тогда условимся, что $\tau(\alpha) = \Lambda$ и $H_\Omega(\alpha) = (\alpha_1, ..., \alpha_{\lambda(\alpha)})$, и для $1 \leq i \leq \lambda(\alpha)$ справедливо $\tau(\alpha_i) = byte$ (сигнатура состоит из последовательности элементарных сигнатур, сумма длин которых равна размеру сигнатуры α). В то же время возможна ситуация, когда тип сигнатуры α не идентифицирован ($\tau(\alpha) = \Lambda$), но она состоит из последовательности сигнатур известных типов, не совпадающих с элементарными. Например, сигнатура α может состоять только из сигнатур – символьных значений полей или сигнатур – числовых значений полей. При этом в совокупности распределение или взаимосвязь входящих в сигнатуру α признаков будут оставаться неизвестными и, следовательно, тип сигнатуры α не будет идентифицирован. Другой вариант: тип сигнатуры α идентифицирован ($\alpha \in \Omega_t$), а множество $H_\Omega(\alpha)$ – нет. Например, сигнатура α представляет собой заголовок пакета, тогда как внутренние сигнатуры, определяющие поля заголовка пакета, не распознаны.

Таким образом, на пространствах сигнатур и типов сигнатур заданы отношения частичного порядка, что позволяет для удобства восприятия представить иерархию сигнатур или иерархии типов сигнатур в виде ориентированного графа. Вершинами графа будут сигнатуры или их типы, а ребра соответствуют отношению частичного порядка между вершинами.

Во многих случаях при идентификации признаков воздействий не анализируются семантика и (или) синтаксис структуры сигнатур непосредственно в признаках (пакетах, кадрах, потоках данных), обрабатываемых ОП [5, 6]. Например, предполагается, что конструкция сигнатуры (последовательности соответствующих им байт) не может искажаться или «склеиваться». В рассматриваемом методе АИПП СПД осуществляется анализ признаков воздействий на ОП с использованием структурных и семантических особенностей сигнатур признаков. Таким образом, целесообразно выполнить анализ сигнатур с учетом возможного синтаксиса, т.е. места их расположения в признаках.

Установим сигнатуры $\alpha, \beta \in \Omega$ такие, что $\alpha < \beta$. Представлением сигнатуры α в сигнатуре β назовем часть сигнатуры последовательности $\nu(\beta)$, соответствующую значению сигнатуры $\nu(\alpha)$. Тогда зададим отображающую функцию $\psi : \{1, ..., \lambda(\alpha)\} \to \{1, ..., \lambda(\beta)\}$ (которую назовем функцией отображения сигнатуры α в сигнатуру β), в которой для любых значений сигнатур $\nu(\alpha) = (\alpha_1, ..., \alpha_{\lambda(\alpha)})$ и $\nu(\beta) = (\beta_1, ..., \beta_{\lambda(\beta)})$ справедливы равенства $\alpha_i = b_{\psi(i)}$, где $1 \leq i \leq \lambda(\alpha)$.

Для сигнатур $\alpha, \beta \in \Omega$, где $\alpha < \beta$, элементы набора $\nu(\beta)$, соответствуют представлению сигнатуры α в сигнатуре β, находятся в произвольном порядке и между ними содержатся элементы, не свойственные значению $\nu(\alpha)$ сигнатуры α.

В ОП, как правило, сигнатуры базовых типов наиболее распространены [4–6]. В связи с этим разработчики

протоколов стремятся унифицировать представление этих сигнатур в признаках, что облегчает разработку идентификаторов протоколов СПД базовых типов. Таким образом, используем следующее утверждение.

Установим сигнатуры $\alpha,\beta \in \Omega$ такие, что $\alpha < \beta$ и сигнатура α имеет базовый тип ($\tau(\alpha) \in T_0$). Тогда функция ψ инициации сигнатуры α в сигнатуру β обладает следующим свойством: для $0 \le i \le \lambda(\alpha)$ выполняется условие $\psi(i+1) = \psi(i) + 1 = \psi(1) + i$.

Таким образом, описана модель идентификации сигнатур, содержащихся в признаках (заголовках пакетов и значениях полей), обрабатываемых ОП, которая определяет формализованное описание сигнатур признаков с позиций их дальнейшей оценки. В итоге полученные соотношения будут использованы на этапе генерации тестовых наборов исходных сигнатур.

Модель протокола СПД, формализующая процесс оценки его защищенности

Разработанный авторами метод АИПП СПД относится к категории динамических [1], то есть он позволяет осуществлять идентификацию признаков в ОП на этапе активного применения протокола как процесса. Следующим шагом работы будет описание модели процесса ОП (протокола СПД), которую в дальнейшем будем использовать для оценки защищенности в ходе применения разработанного метода АИПП СПД.

Установим, что для процесса ОП, реализующего обмен данными, эквивалентен автомат $P\langle S, I, p, s_0\rangle$, где S – множество всех состояний протокола; $I \subset \Omega^*$ – множество входных сигнатур признаков протокола; $p : I \times S \to S$ – функция переходов протокола из состояния в состояние; $s_0 \in S$ – начальное состояние ОП.

Функция перехода p задается кортежем исполняемых операций, которые реализуют процесс P ОП.

Для каждого ОП на множестве состояний S задается пространство аварийных состояний $E \subset S \setminus \{s_0\}$. При этом по определению для каждого аварийного состояния $e \in E$ ОП справедливо утверждение – для любых входных сигнатур $i \in I$ выполняется равенство $p(i,e) = e$ (протокол, перейдя в аварийное состояние, не может перейти ни в какое другое состояние).

Если протокол переходит в аварийное состояние, условимся, что процесс P завершился аварийно.

Таким образом, для любого ОП множество аварийных состояний E можно считать всегда известным, так как оно задается средой СПД, в которой он функционирует. Процесс P завершается аварийно, если выполнение очередной протокольной инструкции «фактически» невозможно (например, нет доступа к порту, невозможно идентифицировать значение поля пакета или значение поля кадра не соответствует разрешенному и т.п.) [4–9].

Эталонные входные сигнатуры – это входные сигнатуры, которые созданы на этапе корректного функционирования ОП, и их обработка по определению не приводит к его аварийному завершению. Если эталонные входные сигнатуры формируют логическую последовательность и являются пакетом, который подается на вход ОП, то это – эталонный пакет.

Учитывая ранее изложенное ($P\langle S, I, p, s_0\rangle$ – процесс ОП), будем говорить, что в протоколе P содержится протокольная ошибка, если существуют входные сигнатуры $i \in I$ такие, что справедливо условие $p(i, s_0) \in E$. Протокольная ошибка – это ошибка в ОП, которая приводит к его аварийному завершению при надлежащих входных сигнатурах.

Исходные данные:

$I_S \subset I$ – множество эталонных входных сигнатур для ОП, при этом для любых эталонных входных сигнатур $i \in I_S$ и любого состояния $s \in S$ справедливо $p(i,s) \in E$;

$I_F \subset I$ – множество входных сигнатур, при обработке которых происходит аварийное завершение ОП $P\langle S, I, p, s_0\rangle$, при этом выполняется условие $I_F \cap I_S = \emptyset$;

$(e, i_1, ..., i_n)$ – траектория протокольной ошибки ОП; при этом $e \in E$ и $(i_1, ..., i_n)$ – последовательность входных сигнатур длины $n \ge 1$ (размер поля заголовка пакета), подаваемых на вход ОП $P\langle S, I, p, s_0\rangle$, где для $1 \le k \le n$ справедливо условие $i_k \in I$ и выполняется равенство $p\left(i_n, p\left(i_{n-1}, ..., p\left(i_2, p\left(i_1, s_0\right)\right)...\right)\right) = e$;

F – множество всех траекторий протокольных ошибок ОП.

В дальнейшем условимся, что в любом ОП существуют протокольные ошибки. Для процесса $P\langle S, I, p, s_0\rangle$ любого ОП существуют входные сигнатуры $(i_1, ..., i_n)$, где $n \ge 1$, и справедливо условие $p((i_1, ..., i_n), s_0) \in E$ (при обработке входных сигнатур $(i_1, ..., i_n)$ протокол $P\langle S, I, p, s_0\rangle$ завершается аварийно). При этом $(p((i_1, ..., i_n), s_0), i_1, ..., i_n) \in F$ является траекторией протокольной ошибки процесса $P\langle S, I, p, s_0\rangle$.

Таким образом, описанная модель позволяет формализовать процесс оценки протокола с позиций идентификации сигнатур признаков, что в дальнейшем используется при разработке СПО для оценки защищенности протоколов СПД.

Метод автоматизированной идентификации признаков протоколов

Дальнейшим этапом разработки метода АИПП СПД является его программная реализация, которая позволит существенно сократить время оценки защищенности протоколов и выполнить необходимые требования по оперативности, своевременности и достоверности. Разработанное специализированное программное обеспечение (СПО) метода АИПП СПД, реализующее следующие функции [10, 11]:

– формирование входных сигнатур для ОП;

– проведение оценки (тестирование ОП – инициализация ОП и передача на его вход подготовленных сигнатур признаков);

– идентификация и регистрация аномалий ОП, возникших в результате обработки входных сигнатур ОП.

В качестве объекта исследования выступает ОП. Суть метода АИПП СПД заключается в выполнении следующих шагов.

Шаг 1. Проводится анализ доступной информации об ОП, т.е.:

– документации (RFC, документация разработчиков и т.д.);

– спецификаций легитимных сигнатур обрабатываемых признаков ОП;

– фрагментов пакетов или кадров (т.е. в среде, где реализуется ОП), которые используют идентификаторы сигнатур, содержащихся во входных признаках ОП.

В результате выполнения шага 1 для ОП:

– описываются множества сигнатур Ω, известных сигнатур Ω_n, типов сигнатур T, типов известных сигнатур Ω_t и идентификаторов сигнатур Θ, функции λ, ν, τ, ω, H_T и H_Ω;

– задается, как в СПД ОП может быть активизирован, т. е. запущен ОП $P\langle S, I, p, s_0\rangle$, и как ему могут сообщаться входные сигнатуры (множество $I \subset \Omega^*$); устанавливается множество входных сигнатур I и их вид (пакеты или их последовательность, поля пакетов и их значения, длины и семантика сигнатур).

Шаг 2. Исходя из результатов выполнения шага 1 выбираются сигнатуры, для которых является необходимым восстановление и анализ их бинарного кода. Выбор может быть осуществлен на основе:

– предположений о наличии протокольных ошибок в бинарных кодах выбранных сигнатур;

– наличия вычислительных ресурсов для проведения детального анализа (идентификация, распознавание по структуре и по семантике признака) бинарного кода сигнатур в ОП, позволяющих выполнить эти процедуры за доступное время.

Для выбранных признаков осуществляется распознавание и анализ бинарного кода, в результате сигнатуры относятся к категории идентифицированных. Кроме того, уточняются значения множеств Ω, Ω_k, T, Ω_t и Θ, функций λ, ν, τ, ω, H_T и H_Ω.

Шаг 3. Учитывая результаты выполнения шагов 1 и 2 выбирают неизвестные сигнатуры (сигнатуры $\alpha \in \Omega_n$), для которых определен тип (сигнатуры из множества Ω_t^*) или распознавание их типа является не целесообразным. Подбор осуществляется с учетом:

– наличия вычислительных мощностей и других ресурсов;

– предположений о наличии протокольных ошибок в идентификаторах сигнатур выбранных типов;

– предположений о возможности идентификации сигнатур за доступное время.

Для отобранных сигнатур осуществляется идентификация их типа. Таким образом, уточняются значения множеств T и Ω_t, функций λ, ν, τ, H_T и H_Ω.

Шаг 4. В СПО выполняется автоматизированная идентификация на предмет наличия протокольных ошибок. Она подразумевает запуск процесса $P\langle S, I, p, s_0\rangle$ ОП (каждый запуск процесса назовем этапом). При выполнении одного этапа процессу $P\langle S, I, p, s_0\rangle$ сообщаются входные сигнатуры, сформированные разработанными процедурами формирования входных сигнатур (ПФВС) из эталонных входных сигнатур путем модификации их значений из множества Ω_t. Если в результате обработки входных сигнатур ОП завершился аварийно, то управление возвращается СПО, которая фиксирует состояние процесса $P\langle S, I, p, s_0\rangle$ и входные сигнатуры, обработка которых стала причиной аварийного завершения. При этом СПО восстанавливает корректность ОП для дальнейшего тестирования.

Шаг 5. После выполнения шагов 1, 2 и 3 выбираются сигнатуры, тип которых не определен и которые целесообразно считать элементарными (тип таких сигнатур равен *byte*). Выбор сигнатур осуществляется с учетом различных факторов:

– предположения о наличии протокольных ошибок в сигнатурах;

– наличия вычислительных ресурсов для проведения за доступное время анализа ОП при обработке входных сигнатур, в которых модифицируются значения сигнатур.

Кроме того, на шаге 5 задается множество сигнатур Ω_b.

Шаг 6. В СПО осуществляет оценка достоверности ОП. При запуске одного этапа ОП $P\langle S, I, p, s_0\rangle$ сообщаются сигнатуры, сформированные ПФВС из эталонных входных сигнатур путем модификации значения сигнатур из множества Ω_b (множества элементарных сигнатур). Если в результате обработки входных сигнатур ОП завершился, то управление возвращается СПО, которое фиксирует состояние процесса $P\langle S, I, p, s_0\rangle$ и входные сигнатуры, обработка которых стала причиной аварийного завершения. При этом СПО восстанавливает корректность процесса $P\langle S, I, p, s_0\rangle$ для дальнейшей оценки.

Таким образом, шаг 1 выполняется для стартовой оценки ОП, а время его выполнения не учитывается при оценке эффективности предлагаемого метода АИПП СПД. Кроме того, выбор сигнатур, которые считаются элементарными, на шаге 5 не требует затрат времени, так как осуществляется случайно из неизвестных сигнатур входных признаков с неопределенным типом.

Для шага 2 предполагается поиск ошибок в сигнатурах сложной структуры с применением «классических» подходов (модификация и отслеживание). Однако получение всех исходных бинарных кодов сигнатур позволит решить задачу более качественно (например, на основе анализа пакетов). С другой стороны, время, необходимое для восстановления исходных бинарных кодов, может оказаться неприемлемым. Обращаем внимание, что шаг 2 выполняется параллельно с шагами 3–6.

Предполагается, что оценка корректности метода АИПП СПД учитывает следующие параметры:

– среднее время обнаружения протокольной ошибки;

– средний объем бинарного кода сигнатур;

– среднее число протокольных ошибок, выявленных на шаге 2, 4 и 6;

– среднее время проведения одного теста на шаге 4 и 6;

– относительное число аварийных окончаний, являющееся усредненным отношением числа аварийно завершившихся этапов к общему числу проведенных этапов на шаге 4 и 6.

Для оценки эффективности предлагаемого метода АИПП СПД используем подход на основе оценки степени адекватности системы к выполнению возложенных на нее задач. Основной особенностью таких систем является то, что при выходе из строя отдельных элементов система не выходит из строя полностью, а продолжает выполнять возложенные на нее функции в несколько меньшем объеме. Поэтому для количественной оценки эффективности таких систем используется некий обобщенный показатель, называемый критерием эффективности – $K_э$. Этот критерий характеризует некоторый средний успех, который достигается в результате функционирования системы. В соответствии с этой особенностью для оценки эффективности метода АИПП СПД будем использовать следующую методику.

1-й этап – анализ назначения, задач и условий функционирования системы. На этом этапе конкретизируется назначение и задачи системы с учетом особенностей выполнения технологического цикла. Этот этап является подготовительным.

2-й этап – выбор критерия эффективности. Для метода АИПП СПД наиболее целесообразным критерием эффективности можно рассматривать математическое ожидание числа задач, решаемых в заданное время:

$$K_э = M[S], \tag{1}$$

где S – число задач, выполняемых в течение заданного времени Δt.

3-й этап – составление структурно-функциональной схемы и расчет надежностных показателей элементов системы. На этом этапе система разбивается на ряд элементов, которые могут находиться в одном из двух возможных состояний: «работа» и «отказ». Для метода АИПП СПД в качестве таких элементов выбираются задачи системы, протоколы и уровни СПД. В результате такого разбиения удается построить структурно-функциональную схему, которая предоставляется в виде ориентированного графа и матриц.

Надежностные характеристики элементов системы задаются в виде вероятности «работы» задач метода АИПП СПД и протоколов и уровней СПД в течение заданного времени Δt. Конкретные значения этих показателей вычисляются с учетом надежностных характеристик аппаратуры, защищенности протоколов ИТКС ОЗУ и характера протокольных воздействий. В результате получаем последовательность вероятностей $p_k\,(k = 1,2,...,m)$. Здесь p_k – вероятность «работы» задачи метода АИПП СПД с индексом k, а m – общее число протоколов и уровней СПД.

4-й этап – определение возможных состояний метода АИПП СПД и показателей эффективности каждого состояния. Если система включает m протоколов, каждый из них может находиться в одном из двух возможных состояний, то общее число всех возможных состояний системы будет равно:

$$N = 2^m. \tag{2}$$

Если обозначить через X_r – состояние, при котором первые r протоколов системы находятся в состоянии «работа», а остальные $m - r$ протоколов – в состоянии отказа, то вероятность такого состояния определиться по формуле:

$$P(X_r) = \prod_{k=1}^{r} p_k \prod_{k=r+1}^{m} (1 - p_k). \tag{3}$$

Используя [5], можно вычислить вероятности всех возможных состояний метода АИПП СПД. В каждом из них СПД способна решить определенное число задач защиты. Таким образом, каждому состоянию системы $x_r\,(r = 1,2,...,N)$ можно поставить в соответствие вероятность такого состояния $P(x_r)$ и число решаемых задач S_r в этом состоянии.

5-й этап – расчет критерия эффективности метода АИПП СПД. Общий критерий эффективности определяется как математическое ожидание числа задач защиты, которые могут быть решены в течение заданного времени Δt. Вероятностью перехода из одного состояния в другое в течение времени Δt пренебрегаем. Тогда

$$K_э = M[S] = \sum_{r=1}^{N} P(X_r) S_r. \tag{4}$$

Кажущаяся простота формулы обманчива. При достаточно большом числе протокольных воздействий на СПД расчет становится достаточно трудным. Поэтому на практике ищут способы упрощения расчетов. Например, при симметричной структуре и одинаковых характеристиках протокольных воздействий достаточно произвести расчет для части СПД. При вычислениях, как правило, исключают из рассмотрения маловероятные состояния метода АИПП СПД.

Задача расчета критерия эффективности может быть решена методом статистического моделирования функционирования метода АИПП СПД.

Рассчитанный критерий эффективности используется при проверке соответствия реальной системы заданным тактико-техническим требованиям или при сравнительной оценке проектов различных вариантов метода АИПП СПД. Задача расчета критерия эффек-

тивности может быть решена методом статистического моделирования функционирования СПД. В качестве примера расчеты коэффициента эффективности представлены на рис. 1.

Таким образом, представленный выше метод позволяет реализовать процедуру автоматизированной оценки защищенности протоколов СПД на предмет наличия признаков воздействий. Указанная процедура обеспечивает существенное уменьшение времени на реализацию процесса оценки защищенности. Одним из важных этапов представленного метода является процедура формирования входных сигнатур для обеспечения оценки защищенности протоколов СПД от различных воздействий. Одним из важных этапов метода АИПП СПД является процедура формирования входных сигнатур (ПФВС), которая в своей работе использует модель идентификации сигнатур.

Процедуры формирования входных сигнатур

Применение разработанного метода выявило определенную взаимосвязь между параметром $K_э$ и применяемыми при тестировании на 4 и 6 шаге ПФВС. В процессе исследования были разработаны стандартизированные (типовые) ПФВС, применимые к большинству ОП. Тем не менее для повышения значения параметра $K_э$ при анализе конкретного протокола требуется разработка (или модификация существующих) индивидуальных ПФВС.

ПФВС должны сформировать исходные сигнатуры, которые во время оценки защищенности протокола будут приняты к обработке.

На шаге 4 и 6 генерируются наборы входных сигнатур для запуска процесса ОП. Входные сигнатуры формируются ПФВС из эталонных входных сигнатур путем модификации значения сигнатур из множеств Ω_t и Ω_b.

В СПО для оценки и идентификации применяются ПФВС двух видов:

– статические ПФВС – используют эвристические правила;

– динамические ПФВС – учитывают параметры входного трафика.

Для генерации статических ПФВС используются различные эвристические правила преобразования сигнатур из множеств Ω_t и Ω_b (шаг 4 и 6). Если сигнатуры, содержащиеся в множествах Ω_t и Ω_b, соответствуют входным эталонным данным, то они преобразуются ПФВС.

Для повышения эффективности метода необходимо сочетать оценку и динамический анализ с учетом графа процесса ОП. Непосредственно перед выполнением шагов 4 и 6 для каждого эталонного файла выполняются следующие действия:

– собирается граф оценки на эталонном протоколе;

– уточняются значения множеств T, Ω_t и Ω_b, функций λ, ν, τ, H_T и H_Ω;

– выбирается статическая ПФВС, наиболее подходящая для проведения очередного теста.

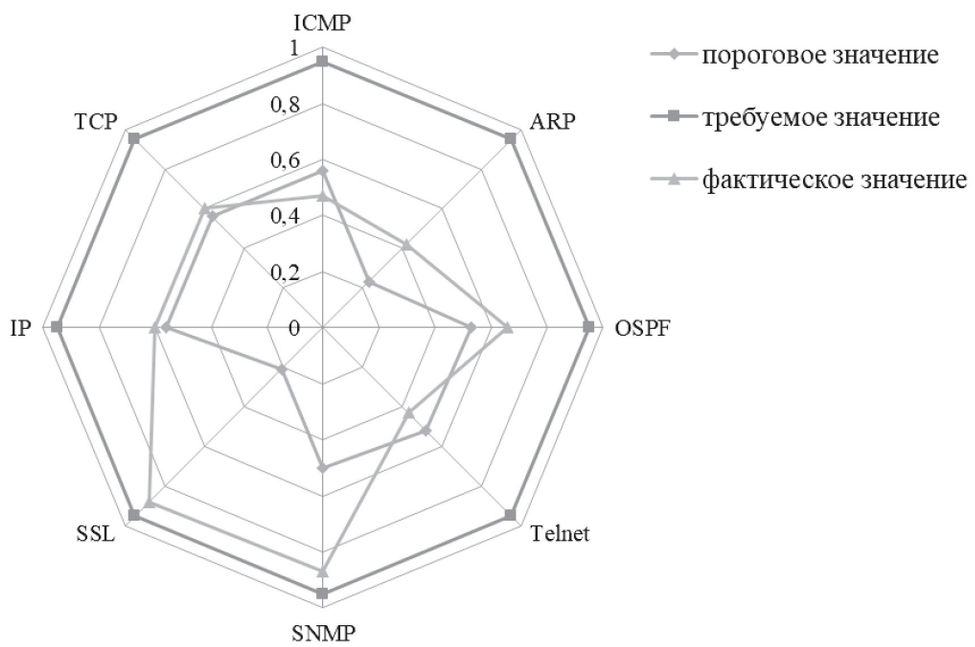

Рис. 1. Оценка эффективности СПЗ ИТКС ОЗУ с помощью графического способа представления результатов

На основе графа оценки эталонного протокола в СПО, которое реализует метод, введены вспомогательные механизмы – это механизм оценки охвата бинарного кода, задействованного при обработке входных сигнатур, и механизм распределения идентификаторов сигнатур за счет сравнения ребер графов. Таким образом, представленная совокупность процедур позволяет выполнить оценку протокола на предмет его достоверности, что в итоге может служить выводом по оценке защищенности протокола СПД.

Заключение

Разработанный метод применяется для оценки защищенности протоколов в СПД различного состава, а также для различных аппаратных и программных платформ. Эксперименты показали хорошие результаты для протоколов, содержащих сигнатуры сложных форматов.

Предлагаемое СПО работает в автоматическом режиме. В качестве результатов работы СПО предоставляет множество сигнатур идентификационных признаков воздействий, вызывающих аварийное завершение протоколов СПД. Таким образом, может быть повышена защищенность разрабатываемых протоколов.

В настоящий момент проводятся исследования по оценке результативности метода с целью анализа его возможностей и расширения границ применения.

Литература

1. Макаров, А. Н. Метод автоматизированного поиска программных ошибок в алгоритмах обработки сложноструктурированных данных / А.Н. Макаров // Прикладная дискретная математика. – 2009. – № 3 (5). – С. 117–127.

2. Attariyan, M. Automating configuration troubleshooting with dynamic information flow analysis / M. Attariyan, J. Flinn // OSDI'10: Proceedings of the 9th USENIX conference on Operating systems design and implementation, 2010. – P. 237–250.

3. Van der Aalst, W. M. P. Workflow mining: Discovering process models from event logs / W.M.P. van der Aalst, T. Weijters, L. Maruster // IEEE Transactions on Knowledge and Data Engineering. – 2004. – Vol. 16, No. 9. – P. 1128–1142.

4. Медведев, Н. В. Применение метода статического сигнатурного анализа для выявления дефектов безопасности веб-приложений / Н.В. Медведев, А.С. Марков, А.А. Фадин // Наука и образование. – 2012. – № 9. – С. 297–314.

5. ГОСТ Р 51188–98. Защита информации. Испытания программных средств на наличие компьютерных вирусов. – Введ. 1999–07–01. – М.: Изд–во стандартов, 1998. – 12 с.

6. Родионов, А. С. Совершенствование методов защиты информации от несанкционированного доступа / А.С. Родионов, В.И. Белянин, А.А. Горбунов // НБИ ТЕХНОЛОГИИ. – 2018. – Т. 12, № 2. – С. 39–43.

7. Анализ свойств вероятностных моментов двоичных последовательностей для задач моделирования стохастических объектов / В.М. Кузнецов [и др.] // Вестник Самарского государственного технического университета. Серия «Технические науки». – 2020. – № 1 (65). – С. 34–48.

8. Московченко, В. М. Робототехническая система анализа кибербезопасности информационных систем и сетей связи / В.М. Московченко, М.А. Гудков, О.С. Лаута // НБИ ТЕХНОЛОГИИ. – 2018. – Т. 12, № 2. – С. 30–38.

9. Дементьев, В. Е. Методика оценки информативности признаков протоколов информационно-телекоммуникационных сетей / В.Е. Дементьев // Технологии и средства связи. – 2016. – № 3 (114). – С. 42–45.

10. Пат. № 2531878 Российская Федерация, МПК G06F11/22. Способ обнаружения компьютерных атак в ИТКС / В.Е. Дементьев [и др.]; патентообладатель Военная академия связи имени Маршала Советского Союза С.М. Буденного Министерства обороны Российской Федерации; заявл. 13.08.2013; опубл. 27.10.2014, Бюл. № 30. – 22 с.

11. Интеллектуальная система обнаружения атак с использованием нейронных сетей / С.В. Костарев. Свидетельство о регистрации программы для ЭВМ RU 2017663191, 27.11.2017. Заявка № 2017616769 от 10.07.2017.