МЕТРИКИ НА ДЕРЕВЬЯХ АТАК, СОГЛАСОВАННЫЕ С МОДУЛЬНОЙ КОМПОЗИЦИЕЙ

Волкова Е. С.[1], Гисин В. Б.[2]

DOI: 10.21681/2311-3456-2024-3-14-22

Цель исследования: *представить общую схему, в рамках которой могут быть сформулированы и вычислены метрики деревьев атак, содержащих гейты конъюнкции, дизъюнкции и секвенциальной конъюнкции.*

Методы исследования: *логико-математический анализ, линейная логика, аппарат теории категорий.*

Полученные результаты: *предложен подход к построению метрик на динамических деревьях атак, основанный на представлении метрики алгеброй над операдой деревьев атак с модульной композицией. Показано, что метрики, вычисляемые методом от листьев к корню, согласуются с модульной композицией. Наличие в дереве атаки узлов секвенциальной конъюнкции индуцирует на множестве терминальных вершин структуру направленного графа. Если этот граф ацикличен, метрики, согласованные с модульной композицией, имеют однозначную интерпретацию. При наличии на графе циклов, однозначность интерпретации обусловлена содержательными свойствами атомарных элементов атаки. В статье показано, что содержательные свойства атомарных элементов могут быть представлены соответствующими тождествами в алгебре термов. Для этого введено понятие дизъюнктивной нормальной формы динамического дерева атаки и показано, что любое дерево может быть представлено в такой форме преобразованиями, использующими только базовые тождества. Научная новизна полученных результатов состоит в применении аппарата операд для определения метрик на динамических деревьях атак.*

Ключевые слова: *дерево атак, секвенциальная конъюнкция, дизъюнктивная нормальная форма, линейная логика, модулярная категория, операда, функтор.*

COHERENT METRICS ON ATTACK TREES

Volkova E. S.[3], Gisin V. B.[4]

The purpose of research: *to present a framework within which metrics of attack trees containing conjunction, disjunction and sequential conjunction gates can be developed and calculated.*

Methods: *mathematical logic, linear logic, machinery of the category theory*

Results: *An approach to the construction of metrics on dynamic attack trees is proposed. A metric is considered as an algebra over the operad of attack trees with modular composition. Such metrics are called consistent with the modular composition. It is shown that the bottom-up calculated metrics are consistent with the modular composition. The presence of sequential conjunction nodes in the attack tree generates a directed graph on the set of the terminal vertices. If this graph is acyclic, a metric consistent with the modular composition have an unambiguous interpretation. If there are cycles on the graph, the unambiguity of interpretation is due to the substantial properties of the basic attack steps. The paper shows that the meaningful properties of atomic elements can be represented by equations in the algebra of terms. For this purpose, the concept of a disjunctive normal form of a dynamic attack tree is introduced and it is shown that any tree can be represented in this form by transformations using only basic identities. The scientific novelty of the results obtained consists in the application of operads to determine metrics on dynamic attack trees.*

Keywords: *attack tree, sequential conjunction, disjunctive normal form, linear logic, modular category, operad, functor.*

1 Волкова Елена Сергеевна, к.ф.-м.н., доцент, Финансовый университет при Правительстве Российской Федерации, Москва, Россия. E mail: evolkova@fa.ru
2 Гисин Владимир Борисович, к.ф.-м.н., профессор, Финансовый университет при Правительстве Российской Федерации, Москва, Россия. E mail: vgisin@fa.ru

3 Elena S. Volkova, Ph.D., Associate Professor, Financial University under the Government of the Russian Federation, Moscow, Russia. E mail: evolkova@fa.ru
4 Vladimir B. Gisin, Ph.D., Professor, Financial University under the Government of the Russian Federation, Moscow,. Russia. E mail: vgisin@fa.ru

Введение

Несмотря на значительный прогресс в разработке методов и средств обеспечения информационной безопасности, число инцидентов кибербезопасности существенно растет как в абсолютном, так и в относительном выражении. По данным, приведенным в материалах Positive Technologies, в IV квартале 2023 г. число кибератак выросло за год на 19%.

Вряд ли можно сомневаться в том, что дилемма «щита и меча» в сфере информационной безопасности не будет разрешена в ближайшем будущем. С учетом этого не теряет своей актуальности совершенствование и разработка новых методов оценки рисков информационной безопасности. Существует множество методов и моделей, которые были разработаны для проведения оценок безопасности.

При моделировании угроз одними из наиболее широко используемых являются модели, в основу которых положено дерево (граф) атак. Используя дерево атак, можно описать цепочки шагов атаки или уязвимостей, которые могут быть использованы злоумышленником для достижения своих целей.

Дерево атак позволяет проводить эффективный анализ безопасности путем систематической организации различных способов, с помощью которых система может быть атакована. Преимущество подобного подхода заключается в сочетании удобных для пользователя, интуитивно понятных визуальных функций с формальной семантикой и алгоритмами, позволяющими проводить качественный и количественный анализ.

Деревья атак были введены как средство представления профиля атакующего. Для достижения некоторой цели инициатор атаки может действовать в соответствии с подцелями. Существует два способа разделения цели: либо цель состоит из множества подцелей, каждая из которых должна быть достигнута; либо цель может быть достигнута с помощью одной из нескольких альтернативных подцелей. Корневой узел дерева атак представляет цель атакующего, а дочерние узлы каждого узла представляют ее уточнение до подцелей. Первоначально рассматривались дизъюнктивные (OR-узел), либо конъюнктивные (AND-узел) уточнения. Листья дерева атак представляют базовые действия атакующего и называются базовыми действиями (BAS). Дерево атак быстро стало популярным инструментом моделирования для анализа безопасности. В течение последних десятилетий графические подходы привлекли внимание многочисленных экспертов по безопасности и формальным методам и стали самостоятельной исследовательской областью (см. [1, 6, 11]).

Развиваются исследования, направленные на конструирование семантики деревьев атак. Ключевым является вопрос, когда два дерева атак могут рассматриваться как представляющие одну и ту же атаку. Если алгоритм или эксперт по безопасности модифицирует дерево атак, желательно знать семантику, чтобы понимать, инвариантны ли свойства атак относительно этих преобразований. Вообще говоря, семантика зависит от типа вопроса, для разрешения которого используется дерево атак, а вопросы характеризуются доменами атрибутов. Например, для вопросов типа «да – нет» подходит семантика, основанная на классической пропозициональной логике. Более общая семантика, основанная на мультимножествах, охватывает более широкий класс вопросов, связанных с такими атрибутами как «минимальная стоимость атаки» или «максимальный ущерб от атаки».

Однако ограничения OR-AND-формализма, в частности в отношении выражения порядка, в котором выполняются различные этапы атаки, были признаны многими авторами (см., [3, 13]). Для моделирования сценариев безопасности часто требуются конструкции, в которых должны быть четко указаны условия порядка выполнения компонентов атаки. Последовательное уточнение отличается от совместного уточнения, поскольку последнее предполагает, что злоумышленник пытается одновременно и независимо достичь нескольких подцелей. Последовательное уточнение подцелей, как и совместное уточнение, требует, чтобы были достигнуты все подцели. При этом некоторые подцели должны быть достигнуты до того, как могут быть достигнуты другие подцели. Для учета таких сценариев понятие дерева атаки было расширено, в деревьях появились узлы типа SAND (последовательные конъюнктивные уточнения).

Деревья с узлами такого типа называют динамическими. Деревья атак представляют собой в этом случае динамические описания уязвимостей системы, которые эволюционируют под влиянием развития системы и новых представлений о возможностях противника. Со временем будут добавлены новые атаки, а существующие атаки будут более конкретизированы. Эквациональная семантика позволяет только сравнить, являются ли два дерева в точности эквивалентными. Если нужно решать вопрос о том, является ли одно дерево атаки специализацией другого дерева атаки, требуется более гибкий аппарат. Для построения семантики динамических деревьев предложено использовать последовательно-параллельные графы, в которых можно представить не только действия атаки, но и причинно-следственные зависимости между действиями. Семантика тесно связана с используемым доменом атрибутов (см. [3]).

Помимо качественного анализа атак, деревья атак могут быть использованы для количественного

анализа. Обычно это делается путем присвоения оценочных значений базовым действиям и последующего расчета оценок всех остальных вершин дерева атаки вплоть до корневой. Например, каждому BAS может быть присвоено значение затрат, представляющее ресурсы, которые злоумышленник должен потратить для выполнения этого BAS, а итоговая оценка корневой вершины даст показатель минимальной стоимости успешной атаки.

Существует множество других подобных показателей, таких как минимальное время успешной атаки, среднее время компрометации, ущерб от атаки и т.п. В [8] можно найти обзор общих алгоритмов для оценки показателей безопасности, в основе которых дерево атак.

С учетом того, что имеется много различных подходов к анализу безопасности на основе дерева атак, становится актуальной разработка общей структуры, в рамках которой показатели, связанные с деревом атак, могут быть сформулированы и вычислены.

В литературе были предложены подходы к общей формализации AT-метрик (см. [4]). Они предполагают, что метрика принимает значения в полукольце. В разных работах используются разные способы определения показателя атаки в терминах базовых значений, что зачастую приводит к несовместимым определениям одного и того же показателя. Например, показатель минимального времени для динамического дерева атаки с последовательным элементом И имеет в литературе различные определения, которые несовместимы даже для небольших примеров. В частности, существует не одно определение метрик полукольца, а, по крайней мере, три несовместимых.

Несовместимость определений может быть связана с тем, что во многих работах метрика определяется вместе с алгоритмом вычисления. В результате метрики часто определяются таким образом, чтобы соответствовать алгоритму. В то же время подходы, связанные с определением метрики непосредственно на основе семантики, приводят к NP-сложным задачам [6].

Таким образом, можно констатировать, что существует потребность в формальной структуре для метрик, связанных с деревьями атак, которая была бы достаточно универсальной. В [12] предложен общий подход к определению метрик на основе операд. Необходимые и достаточные условия того, что метрика вписывается в идеологию операд, достаточно естественны. Почти все известные по публикациям метрики удовлетворяют этим условиям. В [12] детально проработано определение метрик для деревьев атак с узлами типа AND и OR и намечены пути распространения соответствующих конструкций на так называемые динамические деревья и деревья атаки-защиты. В настоящей статье мы, используя идеи из [12] и аппарат операд, даем общее определение метрик для деревьев, содержащих помимо узлов типа AND и OR также и узлы типа SAND. Заметим, что в [12] используется чрезмерно ограничительное понимание уточнения элементарного действия, которое фактически налагает запрет на использование уточняющих модулей, в которых встречаются атомарные действия, уже использованные при построении дерева атак. Чтобы обойти возникающие трудности при определении модульной композиции и сделать конструкцию достаточно общей, в работе введено понятие разметки дерева атак. Это позволяет технически разнести операцию модульной композиции и оценки дерева.

Операды являются естественным инструментом для формализации древовидных структур. Операция композиции в операдах отражает иерархическую природу деревьев. Совокупность деревьев атак можно естественным образом снабдить структурой операды. Общую идею определения когерентности метрики для деревьев атак можно представить следующим образом. Пусть D область (домен) атрибутов, в которой оцениваются BAS. Обозначим через A_n множество деревьев атаки с n терминальными вершинами, представляющими BAS. При этом два n-дерева считаются эквивалентными, если одно получено из другого путем переименования вершин и ребер Оценка терминальных вершин дерева из A_n (с фиксированной нумерацией терминальных вершин) может быть представлена элементом множества D^n. Само дерево задает отображение $D^n \longrightarrow D$, при котором оценке терминальных вершин сопоставляется оценка корневой вершины (цели атаки). Метрику можно считать согласованной (когерентной), если этим определяется морфизм операды деревьев в операду, порожденную доменом D.

Статья организована следующим образом. В следующих далее двух разделах вводятся необходимые понятия. Далее дается определение метрики, согласованной с модульной композицией, и устанавливается когерентность метрик типа «от листьев к корню». В последнем разделе описан подход к алгебраическому представлению содержательных свойств атомарных действий, позволяющий унифицировать подход, основанный на операдах.

Деревья атак

Дерево атак содержит набор терминальных узлов, структурированных с использованием операторов конъюнкции (AND) и дизъюнкции (OR). Терминальные узлы представляют атомарные действия злоумышленника (BAS). Узел AND (соответственно узел OR) считается исполненным, если исполнены все

дочерние узлы (соответственно исполнен по крайней мере один дочерний узел). Множество таких деревьев мы будем обозначать AT, а дерево атаки из AT называть AT-деревом. Множество AT-деревьев может быть расширено, если в дереве атаки допускаются узлы типа SAND. Узел SAND, считается исполненным, если исполнены все его дочерние узлы, причем в указанном порядке.

Деревья атак, содержащие узлы всех трех типов OR, AND и SAND, называют динамическими деревьями атак. Будем называть такие деревья атак DAT-деревьями.

Если не сделаны специальные оговорки, под деревом атак понимается DAT-дерево.

Более формально, дерево атак это тройка $T = (V, E, type)$, где (V, E) – дерево с корневой вершиной r_T, $type:V \rightarrow \{AND,OR,SAND,BAS\}$ – отображение, сопоставляющее вершинам их тип так, что выполняются следующие условия:

(1) $type(v) = BAS$ тогда и только тогда, когда v – терминальная вершина (лист);

(2) если $type(v) = SAND$, то множество дочерних узлов вершины v упорядочено.

Примечание. Здесь и далее применительно к деревьям атак мы используем термины «вершина» и «узел» как синонимы.

Для натурального числа n обозначим через $[n]$ множество $\{1,2,...,n\}$. Тогда упорядоченность из п. (2) предыдущего определения можно понимать как биекцию $\alpha:[n] \rightarrow ch(v)$, где n – и $ch(v)$ – соответственно число и множество дочерних узлов вершины v, а порядок определен естественным упорядочением множества $[n]$.

Через $Leaf(T)$ будем обозначать множество терминальных узлов дерева T.

Пусть задано множество Ω элементарных событий, считающихся неделимыми. Разметкой дерева атаки T будем называть отображение $\lambda:Leaf(T) \rightarrow \Omega$, а дерево с разметкой будем называть размеченным. Разметка «склеивает» некоторые терминальные узлы, превращая дерево T в направленный ациклический граф. Через $BAS(T)$ обозначим образ отображения разметки λ – множество элементарных событий, которые могут произойти в атаках, представленных деревом T.

Приведем ставший уже каноническим пример. На рис. 1 представлено дерево атак, направленных на кражу денег с помощью банкомата. Злоумышленник должен сначала получить соответствующие учетные данные, а затем снять деньги с банковского счета жертвы. Таким образом, корень дерева на рис. 1 уточняется с помощью метода последовательной конъюнкции (дуга со стрелкой). Чтобы получить необходимые учетные данные, злоумышленник должен украсть карту жертвы и получить соответствующий PIN-код. Порядок, в котором будут получены карта и PIN-код, не имеет значения, таким образом стандартное конъюнктивное уточнение И (простая дуга) было использовано для уточнения узла ‘получить учетные данные’. Чтобы получить PIN-код, у злоумышленника есть два варианта: он может либо провести социальную инженерию жертвы, чтобы убедить ее раскрыть секретные четыре цифры, либо найти «шпаргалку» с написанным на нем PIN-кодом. Поскольку любой из этих вариантов обеспечивает получение PIN-кода, узел «получить PIN-код» был уточнен с использованием дизъюнктивного уточнения ИЛИ (без дуги).

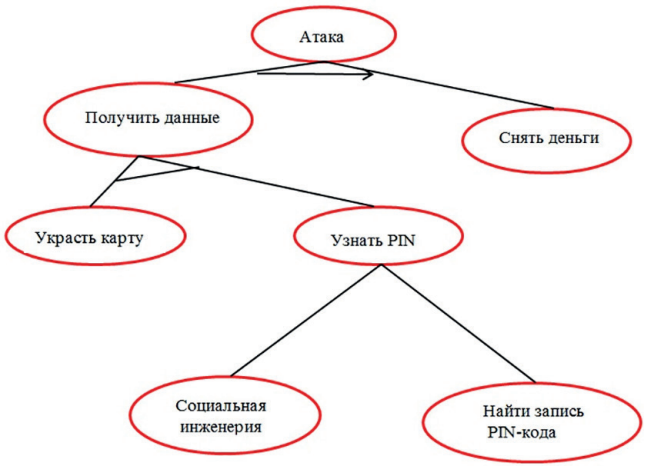

Рис. 1. Дерево атак через банкомат (1)

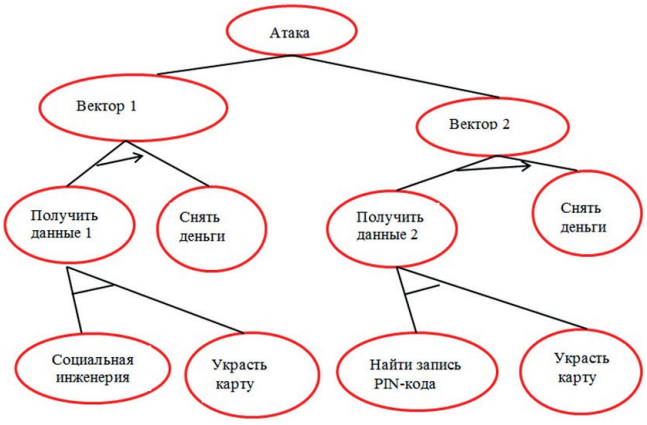

Рис. 2. Дерево атака через банкомат (2)

Дерево на рис. 2 эквивалентно дереву на рис. 1 в том смысле, что представляет те же атаки.

Размеченное дерево атак может быть представлено термом, содержащим имена базовых элементов и операторов. Например, деревья атак на рис. 1 и 2 могут быть представлены соответственно следующими термами

$$t_1 = \mathrm{SAND}(\mathrm{AND}(steal,\mathrm{OR}(eng,pin)),money), \quad (1)$$

$$t_2 = \mathrm{OR}(\mathrm{SAND}(\mathrm{AND}(eng,steal),money), \mathrm{SAND}(\mathrm{AND}(pin,steal),money)), \quad (2)$$

где eng соответствует базовому действию «Социальная инженерия», $steal$ – «Украсть карту», pin – «Найти запись PIN-кода», $money$ – «Снять деньги».

Рассмотрим отношение эквивалентности на множестве термов, порожденное системой определяющих уравнений. Пусть S_n обозначает симметрическую группу, т.е. группу автоморфизмов множества $[n]$. Тогда для любых $k,m \geq 0$ и $l \geq 1$ справедливы тождества (буквы t,u,v с индексом или без индекса служат для обозначения произвольных термов):

(B0) $\mathrm{OR}(t) = t$, $\mathrm{AND}(t) = t$, $\mathrm{SAND}(t) = t$;

(B1) $\mathrm{OR}(t,t,u_1,\ldots,u_n) = \mathrm{OR}(t,u_1,\ldots,u_n)$, $\mathrm{AND}(t,t,u_1,\ldots,u_n) = \mathrm{AND}(t,u_1,\ldots,u_n)$ для любого $n \geq 0$;

(B2) $\mathrm{OR}(t_1,\ldots,t_n,\mathrm{OR}(u_1,\ldots,u_m)) = \mathrm{OR}(t_1,\ldots,t_n,u_1,\ldots,u_m)$, $\mathrm{AND}(t_1,\ldots,t_n,\mathrm{AND}(u_1,\ldots,u_m)) = \mathrm{AND}(t_1,\ldots,t_n,u_1,\ldots,u_m)$, $\mathrm{SAND}(t_1,\ldots,t_n,\mathrm{SAND}(u_1,\ldots,u_m),v_1,\ldots,v_k) = \mathrm{SAND}(t_1,\ldots,t_n,u_1,\ldots,u_m,v_1,\ldots,v_k)$ для любых $n,k \geq 0$, $m \geq 1$;

(B3) $\mathrm{OR}(t_1,\ldots,t_n) = \mathrm{OR}(t_{\sigma(1)},\ldots,t_{\sigma(n)})$, $\mathrm{AND}(t_1,\ldots,t_n) = \mathrm{AND}(t_{\sigma(1)},\ldots,t_{\sigma(n)})$ для любого $n \geq 1$ и любой перестановки $\sigma \in S_n$

(B4) $\mathrm{AND}(t_1,\ldots,t_n,\mathrm{OR}(u_1,\ldots,u_m)) = \mathrm{OR}(\mathrm{AND}(t_1,\ldots,t_n,u_1),\ldots,\mathrm{AND}(t_1,\ldots,t_n,u_m))$, $\mathrm{SAND}(t_1,\ldots,t_n,\mathrm{OR}(u_1,\ldots,u_m),v_1,\ldots,v_k) = \mathrm{OR}(\mathrm{SAND}(t_1,\ldots,t_n,u_1,v_1,\ldots,v_k),\ldots,\mathrm{SAND}(t_1,\ldots,t_n,u_m,v_1,\ldots,v_k))$ для любых $n,k \geq 0$, $m \geq 1$.

Приведенные выше тождества будем называть базовыми.

Используя базовые тождества, несложно показать, что термы (1) и (2) эквивалентны.

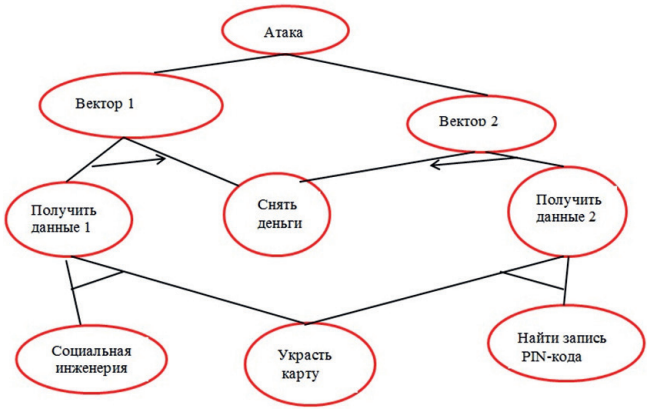

Рис. 3. Граф атака через банкомат

Перечисленные тождества порождают конфлюэнтную систему перезаписи. Соответственно дерево атаки может быть приведено к каноническому виду.

Заметим, что множество терминальных узлов дерева атаки $Leaf(T)$ является мультимножеством с базовым множеством $\mathrm{BAS}(T)$, образованным базовыми элементами атаки.

Более того, узлы типа SAND индуцируют на множестве $Leaf(T)$ отношение частичного порядка. Пусть v – вершина дерева атаки T. Обозначим через $Leaf(v)$ множество терминальных вершин дерева T, в которые из вершины v есть путь по дугам дерева T. Более формально, можно определить $Leaf(v)$ рекурсивно. Если v – терминальная вершина, то $Leaf(v) = \{v\}$; в противном случае $Leaf(v) = \cup_{u \in ch(v)} Leaf(u)$. Очевидно, $Leaf(T)$ является объединением всех множеств $Leaf(v)$. Будем считать, что терминальная вершина a предшествует терминальной вершине b и писать $a \prec b$, если в дереве T существует узел вида $\mathrm{SAND}(u_1,\ldots,u_n)$ такой, что $a \in Leaf(u_k)$, $b \in Leaf(u_m)$ и $k < m$.

Например, для дерева на рис. 1 имеем

$$steal \prec money,\ eng \prec money,\ pin \prec money$$

Оценить успешность атаки, представленной деревом T без узлов типа SAND, можно используя бинарные логические оценки. Пусть $\mathrm{BAS} = \{a_1,\ldots,a_n\}$. Придадим каждому базовому элементу логическое значение «активации» $b_i = I(a_i) \in \boldsymbol{B}$, где $\boldsymbol{B} = \{0,1\}$. Вектор активации $b = (b_i) \in \boldsymbol{B}^n$ приводит к успешной атаке, если оценка по стандартным правилам терма, соответствующего дереву T, дает значение 1.

Для адекватного логического описания деревьев, содержащих последовательную конъюнкцию, приходится использовать более сложную логику. В [6] показано, что адекватным инструментом в этом случае может быть линейная логика, введенная в свое время для описания логики компьютерных программ.

В качестве подходящей шкалы логических значений вместо булевой шкалы $\boldsymbol{B}$ может использоваться шкала $Q = \{0,¼,½,1\}$, содержащая четыре значения. При заданном векторе активации $b = (b_i)$ оценка дерева строится от «листьев к корню» по следующим правилам:

$I[X] = b_i$, если $X \in \mathrm{BAS}$ и $X = a_i$

$I[\mathrm{AND}(X,Y)] = 1$, если $I[X], I[Y] \neq 0$, и $I[\mathrm{AND}(X,Y)] = 0$ в противном случае

$I[\mathrm{OR}(X,Y)] = max(I[X],I[Y])$,

$I[\mathrm{SAND}(X,Y)] = 1$, если $I[X] \geq ½$ и $I[Y] \neq 0$

$I[\mathrm{SAND}(X,Y)] = ¼$, если $I[X] = ¼$ и $I[Y] \neq 0$

$I[\mathrm{SAND}(X,Y)] = 0$, если $I[X] = 0$ и $I[Y] = 0$

Атаку следует признать успешной, если итоговая оценка оказывается положительной.

Если деревья эквивалентны, оценки атак совпадают при любом векторе активации.

Атаки на динамическом дереве атак

Наличие узлов типа SAND делает целесообразным скорректировать понятие атаки (см. [2, 13]).

Пусть T – динамическое дерево атак с разметкой. Для вершины v положим $\mathrm{BAS}(v) = \lambda(Leaf(v))$. Таким образом, $\mathrm{BAS}(v)$ это множество элементарных событий, связанных с вершиной v и/или ее потомками.

Под атакой понимается множество базовых элементов $A \subseteq \mathrm{BAS}(T)$ вместе с отношением частичного порядка $\longrightarrow$. Содержательно $a \longrightarrow b$ можно интерпретировать как то, что a должно быть выполнено перед b. Таким образом, атака $(A,\longrightarrow)$ означает, что выполнены все базовые элементы из A и порядок их выполнения согласуется с $\prec$. Например, для дерева на рис.1

$$(\{pin,steal,money\},\{pin\longrightarrow money,steal\longrightarrow money\}) \quad (1)$$

является атакой. Атакой будет также и

$$(\{pin,steal,money\}, \{pin\longrightarrow money,steal\longrightarrow pin,steal\longrightarrow money\}) \quad (2)$$

В отличие от атаки (1) атака (2) не является минимальной: элемент отношения порядка $steal\longrightarrow pin$ можно удалить.

Отношение порядка $\prec$ должно быть согласовано с отношение порядка на множестве $(Leaf(T)$ в том смысле, что если $\lambda(a),\lambda(b) \in A$ и $\lambda(a)\longrightarrow\lambda(b)$, то $a \prec b$.

Успешность атаки A определим рекурсивно. Атака A успешна, если она достигает корневой вершины r_T. Атака достигает цели (вершины) v если:

(1) $type(v) = \mathrm{BAS}$ и $v \in A$
(2) $type(v) = \mathrm{OR}$ и атака A достигает некоторой вершины $u \in ch(v)$
(3) $type(v) = \mathrm{AND}$ и атака A достигает все вершины $u \in ch(v)$
(4) $type(v) = \mathrm{SAND}$, атака A достигает все вершины из упорядоченного множества $ch(v)=(u_1,\ldots,u_n)$, при этом, если $\lambda(a) \in A \cap \mathrm{BAS}(v_i)$ и $\lambda(b) \in A \cap \mathrm{BAS}(v_{i+1})$, то $\lambda(a)\longrightarrow\lambda(b)$.

В [4] введено понятие правильно сформированного (размеченного) дерева атак. Отношение порядка на множестве $Leaf(T)$ определяет структуру графа на множестве $\mathrm{BAS}(T)$. Считается, что $x,y \in \mathrm{BAS}(T)$ соединены дугой xy, если найдутся такие терминальные вершины $a,b \in Leaf(T)$, что $\lambda(a) = x$, $\lambda(b) = y$ и при этом b непосредственно следует за a относительно упорядочения $\prec$ на множестве $Leaf(T)$. Обозначим этот граф $G(T)$. Дерево T считается правильно сформированным, если граф $G(T)$ ациклический.

Дерево на рис. 1, очевидно, сформировано правильно. Терм

$$t = \mathrm{SAND}(\mathrm{OR}(x,y),\mathrm{OR}(y,z)) \quad (3)$$

дает пример неправильно сформированного дерева T: граф $G(T)$ содержит петлю в вершине y.

В [4] доказано, что пакеты атак на правильно сформированном дереве обладают важным свойством: расширение успешной атаки также успешно. Для неправильно сформированных деревьев это уже не так. Например, для дерева (3) атака

$$A = \{x,z\},\ \prec = \{x\longrightarrow z\}$$

успешна, а атака

$$A = \{x,y,z\},\ \prec = \{x\longrightarrow z\}$$

– нет (в последней для успешности нужно $\prec$ расширить до $\prec = \{x\longrightarrow y,\ x\longrightarrow z\}$).

Полученный в [4] результат позволяет описать семантику правильно сформированных деревьев атак с узлами типа SAND, сопоставляя узлу дерева v пакет минимальных атак, достигающих этот узел. Для получения семантики произвольных деревьев атак этого уже недостаточно. С узлом дерева атак приходится связывать пакет всех атак, достигающих этот узел.

Операды

Важной операцией над деревьями атак является модульная композиция. При модульной композиции базовый элемент атаки дерева T может быть заменен деревом атаки T'. Такая замена используется в том случае, когда нужно детализировать строение элемента атаки, первоначально считавшегося неделимым событием.

Чтобы рассчитать значения тех или иных метрик безопасности, зная дерево атак, можно поступить следующим образом. Фиксируется алгебраическая шкала D с операциями and, or, $sand$, удовлетворяющими соответствующим тождествам. Значения метрики приписываются базовым элементарным действиям, а метрика для всего дерева (представленного термом) рассчитывается от листьев к корню. Актуальным является также подход, при котором тем или иным способом выделяются обеспечивающие успех атаки векторы инициализации, а итоговая оценка получается как оптимальная в том или ином смысле. При таком подходе есть риск получить метрику, которая не согласуется с модульной композицией. Математическим инструментом для точного понимания согласованности служат операды.

Формально операцию модульной композиции можно представить следующим образом. Пусть $T = (V,E,type)$ – дерево атак и $a \in V$ – терминальная вершина, $type(a) = \mathrm{BAS}$. Далее, пусть $T = (V',E',type')$ – также дерево атак с корневой вершиной $r_{T'}$. Без потери общности можно считать, что множества вершин V и V' дизъюнктны. Модульная композиция – это дерево атак $T[T'/a] = (V'',E'',type'')$ такое, что $V'' = (V\backslash\{a\}) \cup V'$; $type''e(v) = type(v)$ при $v \in V$ и $type''(v) = type'(v)$ при $v \in V'$; множество E'' содержит все дуги из E и все дуги из E',

кроме тех, которые заканчиваются или начинаются в вершине a, и дополнено дугами вида (v,r_T) для всех $v \in V$, для которых $a \in ch(v)$.

Нумерацией дерева атаки T будем называть биективное отображение $\mu: Leaf(T) \to [n]$, где n – число терминальных вершин дерева T.

Пусть D – шкала оценок, которые могут быть приписаны вершинам дерева атаки. Метрика Θ, понимаемая в широком смысле, должна приписать оценку дереву атаки, если заданы оценки терминальных вершин. Рассмотрим оценку терминальных вершин $\gamma: Leaf(T) \to D$. Если терминальные вершины занумерованы отображением μ, метрика сопоставляет каждому набору $\bar{d} = (d_1,...,d_n) \in D^n$ значение $\varphi^{\Theta}_{T,\mu}(\bar{d}) \in D$, где d_i – значение оценки γ на элементе из $Leaf(T)$ с номером i, т.е. $d_i = \gamma(\mu^{-1}(i))$. Таким образом, дерево атак T задает отображение $\varphi^{\Theta}_{T,\mu}: D^n \to D$. Предположим, что $\mu': Leaf(T) \to [n]$ – еще одна нумерация. Тогда $\mu' = \sigma \circ \mu$ для некоторой перестановки σ на множестве $[n]$. Вектор оценок при нумерации μ' имеет вид $\bar{d}' = (d'_1,...,d'_n) \in D^n$, где $d'_{\sigma(i)} = d_i$.

Обозначим множество функций из D^n в D через $End_n(D)$. Перестановка σ на множестве $[n]$ определяет биективное отображение $\tau_\sigma: End_n(D) \to End_n(D)$ так, что

$$\tau_\sigma(\varphi)(d_1,...,d_n) = \varphi(d_{\sigma(1)},...,d_{\sigma(n)}). \quad (4)$$

для $\varphi: D^n \to D$.

Тогда

$$\varphi^{\Theta}_{T,\mu}(d_1,...,d_n) = \varphi^{\Theta}_{T,\mu}(d'_{\sigma(1)},...,d'_{\sigma(n)}) = \tau_\sigma(\varphi^{\Theta}_{T,\mu})(d'_1,...,d'_n). \quad (5)$$

Поскольку деревья T и T' различаются только нумерацией базовых событий, одинаковые оценки этих событий должны приводить к одинаковому значению меры Θ, то есть:

$$\varphi^{\Theta}_{T,\mu}(d_1,...,d_n) = \varphi^{\Theta}_{T',\mu'}(d'_1,...,d'_n). \quad (6)$$

Из (5) и (6) следует, что $\tau_\sigma(\varphi^{\Theta}_{T,\mu}) = \varphi^{\Theta}_{T',\mu'}$.

Замечание. Пусть $\Gamma: \Omega \to D$ – оценка элементарных действий (событий) из Ω. Оценку терминальных вершин $\gamma: Leaf(T) \to D$ дерева T будем называть согласованной с разметкой $\lambda: Leaf(T) \to \Omega$, если $\gamma = \Gamma \circ \lambda$.

Трактовка метрики как системы отображений $\varphi^{\Theta}_{T,\mu}$ в сочетании с модульной композицией деревьев естественным образом подводит к определению метрики на деревьях атак с использованием операд.

Общее понятие операды формулируется для объектов моноидальной категории. Мы приведем здесь (и будем использовать) понятие операды множеств.

Операда представляет собой набор $((R_n,\tau_n)_{n\geq 0}, id, *)$, где R_n – множество, $\tau_n: S_n \to \mathrm{Aut}(R_n)$ – гомоморфизм симметрической группы S_n (порядка n) в группу автоморфизмов множества R_n, $id \in R_1$, а $*$ – операция композиции, которая элементу $f \in R_n$ и набору элементов $g_i \in R_{m_i}$, $i=1,...,n$, ставит в соответствие элемент $f*(g_1,...,g_n) \in R_{m_1+...+m_n}$. При этом должны выполняться следующие условия:

(1) $id * f = f*(id,...,id) = f$

(2) пусть $n, m_1,...,m_n, k_1,1,...k_{n,m_n}$ – целые неотрицательные числа, и $f \in R_n$, $g_i \in R_{m_i}$ и $h_{i,j} \in R_{k_{i,j}}$, тогда

$$f*(g_1*(h_{1,1},...,h_{1,m_1}),...,g_n*(h_{n,1},...,h_{n,m_n})) = (f*(g_1,...,g_n)) * (h_1,1,...,h_{n,m_n}))$$

(3) пусть $n, m_1,...,m_n$ – целые неотрицательные числа, и $f \in R_n$, $g_i \in R_{m_i}$, а $\sigma_i \in S_{m_i}$, $i = 1,...,n$, тогда

$$f*(\tau(\sigma_1)(g_1),...,\tau(\sigma_n)(g_n)) = \tau(\sigma_1,...,\sigma_n)(f*(g_1,...,g_n))$$

где $\tau(\sigma_1,...,\sigma_n)$ перестановка порядка $m = m_1+...+m_n$, которая сохраняет на месте последовательные блоки из m_i элементов, переставляя элементы внутри блока, т.е. σ_1 действует на множестве $\{1,...,m_1\}$, перестановка σ_2 – на множестве $\{m_1+1,...,m_2\}$ и т.д.

(4) пусть $n, m_1,...,m_n$ – целые неотрицательные числа, и $f \in R_n$, $g_i \in R_{m_i}$, $i=1,...,n$, а $\sigma \in S_n$, тогда

$$\tau(\sigma)(f) * (g_1,...,g_n) = f*(g_{\sigma(1)},...,g_{\sigma(n)})$$

Пусть $\underline{R} = ((R_n,\tau_n)_{n\geq 0}, id, *)$ и $\underline{R}' = ((R'_n,\tau'_n)_{n\geq 0}, id', *)$ – операды. Морфизмом $F: \underline{R} \to \underline{R}'$ операды $\underline{R}$ в операду $\underline{R}'$ называется набор отображений $F: \underline{R}_n \to \underline{R}'_n$, $n = 0,1,...$, сохраняющих композицию и отображения τ, так что

$$F(f*(g_1,...,g_n)) = F(f) * (F(g_1),...,F(g_n)) \text{ и } F_n \circ \tau_n = \tau'_n \circ F_n$$

(для краткости мы опустили индексы в первом равенстве).

Ключевым примером для нас является операда деревьев атаки.

Для $n \geq 0$ обозначим через AT_n множество деревьев, имеющих n занумерованных терминальных вершин. Формально, элементами множества AT_n являются классы изоморфных деревьев (T,μ), где T дерево атаки с n терминальными вершинами, а $\mu: Leaf(T) \to [n]$ – нумерация терминальных вершин. Вообще, два дерева с нумерацией (T,μ) и (T',μ') считаются изоморфными, если деревья T и T' изоморфны, а нумерации совпадают после изоморфного отождествления.

Допуская некоторую вольность, мы будем говорить о деревьях атаки как элементах множества AT_n, иногда, опуская упоминание о нумерации, если это не ведет к недоразумениям.

Будем считать, что $AT_0 = \emptyset$ и $AT_1 = \{id\}$. Каждой перестановке σ на множестве $[n]$ соответствует биективное отображение $s_\sigma: AT_n \to AT_n$, которое перенумеровывает терминальные вершины: если $(T,\mu) \in AT_n$, то $s_\sigma(T,\mu) = (T,\sigma \circ \mu)$.

Роль операции $*$ играет модульная композиция деревьев. Пусть $(T,\mu) \in AT_n$ и $(T_i,\mu_i) \in AT_{m_i}$, $i = 1,...,n$. Положим

$$T' = T * (T_1,...,T_n) = T[T_1/a_1,...,T_n/a_n] \quad (7)$$

где a_1 – терминальная вершина дерева T, для которой $\mu(a_i) = i$. Тип вершин композиции (7) сохраняется для всех нетерминальных вершин дерева T и всех вершин деревьев $T_1,...,T_n$. Иными словами, для вершины v дерева T' из (7) имеем: $type'(v) = type(v)$, если v – нетерминальная вершина дерева T, и $type'(v) = type_i(v)$, если v – вершина дерева T_i. Таким образом, множество терминальных вершин дерева T' – это объединение терминальных вершин всех деревьев $T_1,...,T_n$. Естественно, $T' \in AT_m$, где $m = m_1 +...+ m_n$.

Нумерация μ' вершина дерева T' определяется следующим образом: если a – терминальная вершина дерева T' и $a \in Leaf(T_i)$, то

$$\mu'(a) = m_1 +...+ m_{i-1} + \mu_i(a)$$

Замечание. Предполагается, что множества вершин разных деревьев дизъюнктны. Чтобы избежать ненужных сложностей, мы говорим о вершинах дерева T' как о вершинах деревьев, из которых они «родом».

Легко видеть, что модульная композиция удовлетворяет всем условиям из определения операды. Операду деревьев атаки мы будем обозначать $\underline{AT}$.

Вторым примером для нас служит операда $\underline{\mathrm{End}}(D)$. Эта операда образована семейством множеств $\mathrm{End}_n(D)$. Отображения τ определены формулой (4), единицей id служит тождественное отображение $1_D{:}D \longrightarrow D$ из $\mathrm{End}_1(D)$, а операция композиции определена композицией отображений. Заметим, что D^0 – одноточечное множество, так что можно считать, что $\mathrm{End}_0(D) = D$.

Метрику со шкалой D будем называть согласованной с модульной композицией, если она задается морфизмом операд $\underline{\mathrm{AT}} \longrightarrow \underline{\mathrm{End}}\,(D)$.

Рассмотрим метрики, когда в качестве шкалы берется множество D, на котором заданы три ассоциативные бинарные операции $\nabla,\Delta,\Diamond$ такие, что ∇ и Δ коммутативны, а Δ и $\Diamond$ дистрибутивны относительно ∇. Пусть T – дерево атак, имеющее n занумерованных терминальных вершин, а $\gamma{:}Leaf(T) \longrightarrow D$ – оценка терминальных вершин в шкале D. Определим рекурсивно $\varphi_T(\gamma) \in D$:

$$\varphi_T(\gamma)(v) = \gamma(v), \text{ если } type(v) = \mathrm{BAS}$$
$$\varphi_{T,\gamma}(v) = \nabla_{u \in ch(v)}\, \varphi_{T,\gamma}(u), \text{ если } type(v) = \mathrm{OR}$$
$$\varphi_{T,\gamma}(v) = \Delta_{u \in ch(v)}\, \varphi_{T,\gamma}(u), \text{ если } type(v) = \mathrm{AND}$$
$$\varphi_{T,\gamma}(v) = \Diamond_{u \in ch(v)}\, \varphi_{T,\gamma}(u), \text{ если } type(v) = \mathrm{SAND}$$

Наконец, $\varphi_T(\gamma) = \varphi_{T,\gamma}(r_T)$ – значение $\varphi_{T,\gamma}$ в корневой вершине дерева T. Про метрику $\gamma \mapsto \varphi_T(\gamma)$ будем говорить, что она получена методом от листьев к корню и является bu-метрикой (от bottom-up).

Соответствие $T \mapsto \varphi_T$ задает отображение $F_n{:}AT_n \longrightarrow \mathrm{End}_n(D)$. Следующая теорема утверждает, что отображения F_n определяют морфизм $\underline{\mathrm{AT}} \longrightarrow \mathrm{End}(D)$.

Теорема. Любая bu-метрика согласована.

Доказательство достаточно очевидно. Проверка того, что необходимые условия выполняются, может быть проведена прямым вычислением.

В качестве примера рассмотрим метрику, определяющую минимальное время атаки. Пусть $D = \boldsymbol{R}_{\geq 0}$ – множество неотрицательных действительных чисел,

$$x \nabla y = min(x,y),\ x \Delta y = max(x,y),\ x \Diamond y = x+y$$

(события, связанные с узлом AND могут быть выполнены параллельно, с узлом SAND – только последовательно).

Тогда, например, для дерева на рис. 2 при нумерации вершин слева направо имеем

$$\varphi_T(\gamma) = min(max(\gamma_1,\gamma_2) + \gamma_3,\ max(\gamma_4,\gamma_5) + \gamma_6)$$

При этом, очевидно, должны выполняться равенства $\gamma_2 = \gamma_5$ и $\gamma_3 = \gamma_6$ поскольку в первом случае обе оценки относятся к одному и тому же элементарному событию «украсть карту», а во втором – «снять деньги».

Тождества в алгебре термов

То, что дерево атак после разметки терминальных вершин фактически превращается в направленный граф без циклов, налагает ограничения на применимость метрики, согласованной с модульной композицией. Если дерево T правильно сформировано, т.е. граф $G(T)$ ациклический, то проблем не возникает. Если же на графе $G(T)$ имеются циклы, bu-оценка может оказаться неадекватной.

Рассмотрим, например, дерево атак T, представленное следующим термом:

$$t = \mathrm{SAND}(\mathrm{AND}(X,Y),\mathrm{AND}(X,Z))$$

В соответствии с bu-оценкой (для минимального времени) получаем:

$$\varphi_T(\gamma) = max(\gamma(X),\gamma(Y)) + max(\gamma(X),\gamma(Z)).$$

При $\gamma(X) > \gamma(Y)$ и $\gamma(X) > \gamma(Z)$ это приводит к $\varphi_T(\gamma) = 2\gamma(X)$. В то же время, если для исполнения Z не требуется повторного исполнения X, то $\varphi_T(\gamma) = \gamma(X) + \gamma(Z)$. Фактически последнее означает, что используется тождество

$$\mathrm{SAND}(\mathrm{AND}(X,Y),\mathrm{AND}(X,Z)) = \mathrm{SAND}(\mathrm{AND}(X,Y),Z). \quad (8)$$

и bu-оценка строится для его правой части.

Это тождество отражает содержательные связи между элементарными событиями. Подобными тождествами может быть представлено «знание» о связи событий.

Будем говорить, что дерево T представлено в дизъюнктивной нормальной форме (ДНФ), если выполняются следующие условия:

(1) корень r_T – единственная вершина типа OR;
(2) если v – вершина типа AND, то все вершины из $ch(v)$ имеют тип SAND;
(3) если v – вершина типа SAND, то все вершины из $ch(v)$ имеют тип AND;

Используя базовые тождества, любое дерево атак можно представить в виде ДНФ. Это утверждение доказывается стандартным образом по индукции. Выполнение условия (1) обеспечивается применением тождеств (B2) и (B4), условий (2) и (3) – тождеств (B2).

Таким образом, можно ограничиться рассмотрением метрик для деревьев, представленных в ДНФ.

Дополнительные тождества в этом случае будут связывать чередующиеся операции AND и SAND, типа тождества (8).

Например, если базовые события действия такие, что однократное исполнение события не требует его повторного исполнения в последующем, то циклы на графе $G(T)$ можно исключить, введя соответствующий набор тождеств. Тождества этого набора могут быть построены по следующей схеме.

Если в формуле имеется терм вида $\mathrm{SAND}(X_1,\ldots,X_n)$ такой, что $X \in Leaf(X_k) \cap Leaf(X_m)$ при $k < m$, то X может быть исключен из всех термов, составляющих X_m.

Например,

$$\mathrm{SAND}(X,\mathrm{AND}(Y,\mathrm{SAND}(X,Z))) =$$
$$= \mathrm{SAND}(X,\mathrm{AND}(Y,\mathrm{SAND}(Z))) = \mathrm{SAND}(\mathrm{AND}(Y,Z)).$$

Таким образом, получающееся многообразие термов служит алгебраическим эквивалентом содержательного свойства базовых действий.

Заключение

В работе предложен подход к определению метрик на динамических деревьях атак. Выделен класс метрик, которые согласуются с модульной композицией деревьев. Это метрики, которые интерпретируются как морфизмы операды деревьев в операду, порожденную измерительной шкалой как объектом соответствующей симметричной моноидальной категории. Показано, что с модульной композицией согласуются метрики, вычисляемые от листьев к корню. Наличие на дереве атак узлов с последовательной конъюнкцией типа SAND может в некоторых случаях привести к оценкам, имеющим неоднозначную интерпретацию. Добиться однозначности можно за счет более полного содержательного анализа базовых элементарных действий. Такой анализ может быть выражен алгебраически как система тождеств в алгебре термов, представляющих деревья атак. Задачей дальнейшего исследования может быть дальнейшее формирование унифицированного подхода к определению метрик на деревьях атак. В частности, описание классов деревьев атак, для которых имеются нетривиальные и практически значимые метрики, согласованные с модульной композицией.

Литература

1. *Agyepong E. Cherdantseva Y., Reinecke P., Burnap P. Challenges and performance metrics for security operations center analysts: a systematic review //Journal of Cyber Security Technology. – 2020. – Т. 4. – №. 3. – С. 125–152.*
2. *Ali A. T., Gruska D. Dynamic attack trees methodology // 2022 Interdisciplinary Research in Technology and Management (IRTM). – IEEE, 2022. – С. 1–9.*
3. *Bossuat A., Kordy B. Evil Twins: Handling Repetitions in Attack–Defense Trees: A Survival Guide //Graphical Models for Security: 4th International Workshop, GraMSec 2017, Santa Barbara, CA, USA, August 21, 2017, Revised Selected Papers 4. – Springer International Publishing, 2018. – С. 17–37.*
4. *Budde C. E., Stoelinga M. Efficient algorithms for quantitative attack tree analysis //2021 IEEE 34th Computer Security Foundations Symposium (CSF). – IEEE, 2021. – С. 1-15.*
5. *Buldas, A., Gadyatskaya, O., Lenin, A., Mauw, S., & Trujillo-Rasua, R. Attribute evaluation on attack trees with incomplete information // Computers & Security. – 2020. – Т. 88. – С. 101630.*
6. *Eades III H., Jiang J., Bryant A. On linear logic, functional programming, and attack trees //Graphical Models for Security: 5th International Workshop, GraMSec 2018, Oxford, UK, July 8, 2018, Revised Selected Papers 5. – Springer International Publishing, 2019. – С. 71–89.*
7. *Федорченко Е. В., Котенко И. В., Федорченко А. В., Новикова Е. С., Саенко И. Б. Оценивание защищенности информационных систем на основе графовой модели эксплойтов //Вопросы кибербезопасности. – 2023. – №. 3. – С. 23–36.*
8. *Konsta, A. M., Lafuente, A. L., Spiga, B., & Dragoni, N. Survey: Automatic generation of attack trees and attack graphs //Computers & Security. – 2024. – Т. 137. – С. 103602.*
9. *Lallie H. S., Debattista K., Bal J. A review of attack graph and attack tree visual syntax in cyber security //Computer Science Review. 2020. Т. 35. С. 100219. https://doi.org/10.1016/j.cosrev.2019.100219*
10. *Lopuhaä-Zwakenberg M., Budde C. E., Stoelinga M. Efficient and Generic Algorithms for Quantitative Attack Tree Analysis // IEEE Transactions on Dependable and Secure Computing. 20(5). 2022. – 4169–4187. DOI: 10.1109/TDSC.2022.3215752*
11. *Lopuhaä-Zwakenberg M., Stoelinga M. Attack time analysis in dynamic attack trees via integer linear programming //International Conference on Software Engineering and Formal Methods. – Cham : Springer Nature Switzerland, 2023. – С. 165–183.*
12. *Lopuhaä-Zwakenberg M. Attack tree metrics are operad algebras //arXiv preprint arXiv:2401.10008. – 2024.*
13. *Wu, Z., Hu, J., Zhang, X., & Ren, W. timeTree: How to Represent Time Sequence in a Threat Tree //2022 IEEE 24th Int Conf on High Performance Computing & Communications; 8th Int Conf on Data Science & Systems; 20th Int Conf on Smart City; 8th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys). – IEEE, 2022. – С. 2373–2378.*
14. *Zeng J., Wu, S., Chen, Y., Zeng, R., & Wu, C.Survey of attack graph analysis methods from the perspective of data and knowledge processing // Security and Communication Networks. 2019. Т. 2019. Article ID 2031063, 16 С., 2019. https://doi.org/10.1155/2019/2031063*