Modelling Security Risk Scenarios Using Subjective Attack Trees

Nasser Al-Hadhrami[0000−0003−2484−7444], Matthew Collinson[0000−0002−0496−2990], and Nir Oren[0000−0002−4854−9014]

University of Aberdeen, AB24 3UE, Aberdeen, UK
{r01nama,matthew.collinson,n.oren}@abdn.ac.uk

Abstract. We propose a novel attack tree model, called a subjective attack tree, aiming to address the limitations of traditional attack trees, which use precise values for likelihoods of security events. In many situations, it is often difficult to elicit accurate probabilities due to lack of knowledge, or insufficient historical data, making the evaluation of risk in existing approaches unreliable. In this paper, we consider the modelling of uncertainty about probabilities, via subjective opinions, resulting in a model taking second-order uncertainty into account. We propose an approach to derive subjective opinions about security events based on two main criteria, namely a vulnerability level and technical difficulty to conduct an attack, using subjective logic. These subjective opinions are then used as input parameters in the proposed model. The propagation method of subjective opinions is also discussed. Our approach is evaluated against traditional attack trees using the Stuxnet self-installation scenario. Our results show that taking uncertainty about probabilities into account during security risk analysis can lead to different outcomes, and therefore different security decisions.

Keywords: Attack trees · Risk analysis · Subjective logic.

1 Introduction

Attack trees (ATs) [19] have been widely used in recent years as an effective model to analyze security of systems against potential cyber-attacks. One important parameter in ATs used to analyse security risk is the likelihood of *successful attacks* (in literature, also referred to as security events). However, several probabilistic ATs [2, 5, 12, 16, 17, 20] use precise values for likelihoods using the probabilistic approach. In many situations, it is difficult to elicit accurate probabilities due to lack of knowledge, or insufficient historical data, making the evaluation of risk in existing approaches unreliable.

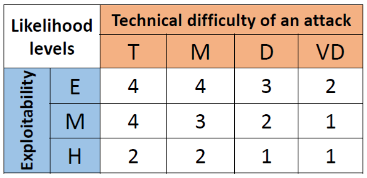
Furthermore, the determination of likelihoods in ATs is not based on a solid foundation based on specific criteria, but rather on a direct assignment of values to ATs leaves. To address this weakness, Abdo [1] proposed the modelling of additional information about security events, e.g., vulnerability information, and that the successful occurrence of attacks is evaluated according to two criteria, namely a vulnerability level (i.e., how easy or hard is to exploit a vulnerability) and technical difficulty to conduct an attack, described by two qualitative scales (see Fig 1) as follows: easy (E) , medium (M), and hard (H), for the vulnerability level, and trivial (T), moderate (M), difficult (D), and very difficult (VD), for the technical difficulty (a detailed description of these two scales can be found in [1]). The final output, representing likelihoods of security input events, is then obtained from combining the qualitative expressions of the two criteria in a form of a matrix as depicted in Fig 1. The work, however, has two major problems. First, it provides only a qualitative evaluation of ATs, and is therefore not suitable for effective decision-making that requires numerical values to make sound decisions. Second, the determination of a vulnerability level and technical difficulty of an attack in a precise manner is often difficult. With continuous emergence of new vulnerabilities— the so called zero-day vulnerabilities— security analysts might be unable to give precise evaluations about their risk levels. In addition, attackers nowadays may have the skills that enable them to conduct cyber-attacks successfully (or discover new attack strategies) even in presence of protected devices and networks with various security technologies. Therefore, it's difficult to precisely evaluate the level of technical difficulty to conduct an attack. Based on such reasons, it is essential to find a way that allows for the modelling of *uncertainty* about the *values* (i.e., the levels) of the two criteria.

In this paper, we address the current limitations of ATs by allowing for *uncertainty* modelling about likelihoods, via *subjective opinions*. In Subjective Logic [9], a subjective opinion represents the probability distribution of a random variable complemented by an *uncertainty* degree about the distribution. Our approach results in a model taking *second-order uncertainty*, i.e., uncertainty about probabilities, into account. We refer to such an AT model as a *Subjective Attack Tree*, abbreviated SAT. We use the evaluation matrix in Fig 1 as one possible way to derive subjective opinions about security events in absence of knowledge or evidence about the evaluation of the two criteria of a vulnerability level and technical difficulty of an attack. Hence, the SAT model (the abstract model in Section 3 and propagation method in Section 5) can be used independently from the evaluation methodology we propose in Section 4 if security analysts prefer to directly assign opinions to the leaves, or if they wish to consider different evaluation methodologies. In comparison to ATs, the SAT model adds a bit more complexity in that it allows also to propagate uncertainty values so that uncertainty about likelihoods of the top events (i.e., root nodes) is also computed.

Explicitly modelling uncertainty degrees about the input parameters in ATs is important as this may lead to different outcomes, e.g., different attack paths prioritization, different enforced sets of countermeasures, different decisions. Apart from such importance, explicitly taking uncertainty about probabilities into account offers a more flexible approach to decision-making process based on factors such as organisations' financial capabilities (budget), risk attitudes, etc. Suppose for example a security analyst is *completely* uncertain about whether an attacker can successfully conduct an attack. In contrast to guessing single probabilities (in absence of knowledge/evidence), our approach allows, for instance, risk-averse security managers to consider the worst-case scenario (pessimistic view) and make decisions so as to protect the system. Others who are risk-seeking, especially those with limited budget, may consider the best-case scenario (optimistic view), and therefore will not need to spend more to protect systems. Decision-making in traditional probabilistic approach leads always to applying strict single decisions under all circumstances.

This work makes the following major contributions. (1) we develop a new model of ATs, called SAT, that takes second-order uncertainty into account. (2) we propose a methodology to derive opinions about security events based on the two criteria discussed in [1] using Subjective Logic. (3) we conduct an experimental evaluation that compares our approach with traditional ATs, demonstrating that the results differ and would lead to different decisions being made.

The rest of the paper is organised as follows. In Section 2, we give an overview of attack trees and discuss some related work. In Section 3, we give an overview of Subjective Logic. In Section 4, we discuss our SAT model, followed by an approach, in Section 5, to evaluate likelihoods of security events using Subjective Logic. In Section 6, we discuss the propagation method of subjective opinions in SATs. In Section 7, we evaluate our approach against traditional ATs, using the Stuxnet attack tree example. Finally, in Section 8, we conclude the paper, discussing prospects for future work.



Likelihood levels		Technical difficulty of an attack			
		T	M	D	VD
Exploitability	E	4	4	3	2
	M	4	3	2	1
	H	2	2	1	1

Fig. 1: The evaluation matrix of security events as proposed in [1].

2 Attack Trees and Related Work

An attack tree (AT) was first introduced in 1999 by Schneier [19] as a tool to analyse and evaluate all possible attack scenarios against complex systems in a structured, hierarchical way. The general idea of ATs is to identify one or more *attack goals* against a system and then break down each goal into sub-goals (or sub-attacks), which in turn can be further broken down into other sub-goals, until reaching a state where sub-attacks cannot be further refined. These final sub-attacks, representing the leaves of an AT, are the basic security events (or action) an attacker can perform, by exploiting existing vulnerabilities, to achieve their overall goal, i.e., the root node of an AT. A refinement from the root node to the leaves can be either conjunctive (via AND node) or disjunctive (via OR node). With AND node, *all* children nodes must be satisfied to complete an attack, while with OR node, *at least one* of the children nodes has to be satisfied.

The values of nodes in a tree can be of different forms, depending on the security attributes or properties need to be analysed. Such values may represent the probability of success of a given attack, the likelihood that an attacker will try a given attack, the impact of an attack, and so on. Earlier works in this field considered attack trees using only one estimated parameter, such as attack probability, cost or feasibility of the attack, skill level required, etc. [13, 14, 19]. Opel [15] considered multi-parameter attack trees (attack trees that study several security attributes of interest), but the actual tree computations in their model still use only one input parameter at a time.

An advanced step towards better understanding the attacker's motivation was made in [3]. The authors considered a multi-parameter attack tree where security properties of interest need to be analysed represent, for examples, gain of the attacker, probability of success, probability of getting caught, and expected penalties.

The above models of ATs have a significant drawback when they come to practical application. The input parameters considered to be precise point estimates based on the probabilistic approach. In [10], the authors addressed this point by suggesting the use of interval values to estimate the input parameters rather than single values. Their approach was basically intended to handle the estimation problem in the multi-parameter AT approach of [3]. While interval values may be a useful method to model the uncertainty about some input parameters, e.g., cost, expected penalties, they are still incapable to model ignorance of or complete uncertainty about likelihoods evaluations of attacks. In addition, specifying lower and upper bounds do not resolve the issue on how these values were precisely determined.

A fuzzy logic approach was employed to model uncertainty in ATs [4]. The approach is based on defining a set of qualitative expressions of likelihoods (e.g., very low, low, high) that describe various levels of likelihoods, and then uses fuzzy numbers to represents experts' judgments on them. The fuzzy logic approach is suitable for applications that involve fuzzy sets, and when there is some difficulty in determining the exact set that a given value should belong to. However, the approach does not model well situations when there is, for instance, a complete uncertainty about the evaluations.

A Bayesian network approach for ATs is explored in [8]. The authors proposed a methodology that translates ATs into Bayesian Networks. The proposed approach can deal with different ATs extensions, and allows the quantitative evaluation of combined attacks modelled as a set of ATs. The Bayesian network approach considers the conditional relations between the nodes, and does not say anything about the values of the leaves (i.e., it employs also the probabilistic approach to assign precise values to the security events).

Our approach differs from all above in that it runs under second-order uncertainty (i.e., uncertainty about probability values) using subjective logic. This allows to better model situations when there is high (or even complete) uncertainty about exact values. Furthermore, subjective logic offers a methodology that easily allows to establish opinions from verbal categories because people often find it difficult to express opinions as numerical values— qualitative verbal categories are intuitively easier [9].

3 Subjective Logic

Subjective logic [9] is a formalism for reasoning under uncertainty that extends probabilistic logic by allowing also for uncertainty degrees to be expressed about probability values. While the idea of probabilistic logic is to combine the strengths of probability calculus and logic, the idea of Subjective Logic is to model uncertainty about the probabilities themselves, making itself a useful tool to reason with argument models in presence of uncertain or incomplete evidence.

Subjective Logic is based on Dempster-Shafer (also called evidence) theory [7], and thus operates on a *frame of discernment*, denoted by Θ, representing the set of possible system states, referred to as atomic, or primitive, system states, only one of which represents the actual system state.

In many scenarios, it is often difficult to determine the actual system state, and it thus makes sense to define non-atomic (or non-primitive) states, consisting of the union of a number of primitive states. The powerset of Θ, denoted by 2^{Θ}, consists of all possible unions of primitive states. A non-primitive state may contain other states within it. These are referred to as substates of the state.

Definition 1. *(Belief Mass Assignment) Given a frame of discernment Θ, we can associate a belief mass assignment $m_{\Theta}(x)$ with each substate $x \in 2^{\Theta}$ such that $m_{\Theta}(x) \geq 0$, $m_{\Theta}(\emptyset) = 0$, and $\sum_{x \in 2^{\Theta}} m_{\Theta}(x) = 1$. For a substate x, $m_{\Theta}(x)$ is its* belief mass.

Subjective logic operates on a 3-dimensional metric called *opinion*. Three classes (types) of opinions are defined, namely *binomial* opinions, *multinomial* opinions, and *hyper* opinions. In this paper, we deal only with binomial opinions.

Definition 2. *(Binomial opinion) Let $X = \{x, \bar{x}\}$ be a state space containing x and its complement $\bar{x}$. A binomial opinion about the truth of state x is the tuple $\omega_x = \langle b_x, d_x, u_x, a_x \rangle$, where b_x is the belief mass in support of x being true, d_x is the belief mass in support of x being false, u_x is the amount of uncommitted belief mass, and a_x is the a priori probability, also called the* base rate*, in the absence of committed belief mass. Further, these components must satisfy $b_x + d_x + u_x = 1$ and $b_x, d_x, u_x, a_x \in [0, 1]$.*

A subjective opinion with $u_x = 0$ is called a *dogmatic opinion*, and corresponds to the classic probability distribution. A dogmatic belief for which $b_x(x) = 1$, for some $x \in \mathbb{X}$, is called an *absolute opinion*. An opinion with $u_x = 1$ is called a *vacuous opinion*. For a given binomial opinion ω_X, the corresponding *projected probability distribution* $\mathbf{P}(x) : x \to [0, 1]$ is determined as

$$\mathbf{P}(x) = b_x + a_x \cdot u_x \tag{1}$$

where $\mathbf{P}(x)$ represents the probability estimation of x which varies from the base rate value, in the case of complete ignorance ($u_x = 1$), to the actual probability in case that $u_x = 0$.

Subjective Logic provides a standard set of logical operators. In this paper we need to deal with only three operators. These are the conjunction (also called multiplication), disjunction (also called co-multiplication), and addition operators.

Definition 3. *(Conjunction Operator) Given two opinions $\omega_x = \langle b_x, d_x, u_x, a_x\rangle$ and $\omega_y = \langle b_y, d_y, u_y, a_y\rangle$ where x and y belong to independent frames of discernment, we compute the conjunction of the two opinions, $\omega_{x\wedge y}$, as*

$$\begin{aligned}
b_{x\wedge y} &= b_x b_y + \frac{(1-a_x)a_y b_x u_y + a_x(1-a_y)u_x b_y}{1-a_x a_y},\\
d_{x\wedge y} &= d_x + d_y - d_x d_y,\\
u_{x\wedge y} &= u_x u_y + \frac{(1-a_y)b_x u_y + (1-a_x)u_x b_y}{1-a_x a_y},\\
a_{x\wedge y} &= a_x a_y.
\end{aligned}$$

By using the symbol $(\cdot)$ to denote this operator, multiplication of opinions can be written as $\omega_{x\wedge y} = \omega_x \cdot \omega_y$.

Definition 4. *(Disjunction Operator) Given two opinions $\omega_x = \langle b_x, d_x, u_x, a_x\rangle$ and $\omega_y = \langle b_y, d_y, u_y, a_y\rangle$ where x and y belong to independent frames of discernment, we compute the disjunction of the two opinions, $\omega_{x\vee y}$, as*

$$\begin{aligned}
b_{x\vee y} &= b_x + b_y - b_x b_y,\\
d_{x\vee y} &= d_x d_y + \frac{a_x(1-a_y)d_x u_y + (1-a_x)a_y u_x d_y}{a_X + a_y - a_x a_y},\\
u_{x\vee y} &= u_x u_y + \frac{a_y d_x u_y + a_x u_x d_y}{a_x + a_y - a_x a_y},\\
a_{x\vee y} &= a_x + a_y - a_x a_y.
\end{aligned}$$

By using the symbol $(\sqcup)$ to denote this operator, co-multiplication of opinions can be written as $\omega_{x\vee y} = \omega_x \sqcup \omega_y$.

Definition 5. *(Addition Operator) Given two opinions $\omega_x = \langle b_x, d_x, u_x, a_x\rangle$ and $\omega_y = \langle b_y, d_y, u_y, a_y\rangle$ where x and y be two disjoint subsets of the same frame X, i.e., $x \cap y = \emptyset$, we compute the addition of the two opinions, $\omega_{x\cap y}$, as*

$$\begin{aligned}
b_{x\cap y} &= b_x + b_y,\\
d_{x\cap y} &= \frac{a_x(d_x - b_y) + a_y(d_y - b_x)}{a_x + a_y},\\
u_{x\cap y} &= \frac{a_x u_x + a_y u_y}{a_x + a_y},\\
a_{x\cap y} &= a_x + a_y.
\end{aligned}$$

By using the symbol $(+)$ to denote this operator, addition of opinions can be written as $\omega_{x \cap y} = \omega_x + \omega_y$.

4 Subjective Attack Trees

In this section, we discuss our approach to model security risk scenarios under second-order uncertainty, using Subjective Attack Trees (SATs).

In SATs, the tree structure is not different from the one in traditional ATs in that it also allows for the decomposition of the main goal of an attacker into sub-goals either conjunctively or disjunctively, except that the input parameters represent subjective opinions rather than probabilities.

Fig 2 shows an example SAT with three possible paths (ways) an attacker can choose to achieve their main goal (MG). These paths begin by the execution of the following security events: (SE_1 and SE_2), SE_3, and (SE_4 and SE_5). Taking the first path with security events SE_1 and SE_2 as an example, the subjective opinions on them, respectively, are denoted by ω_{SE_1} and ω_{SE_2}. The subjective opinion on sub-goal 1 (ω_{SG_1}) is computed from the *conjunction* of ω_{SE_1} and ω_{SE_2}, and the subjective opinion on the main goal (ω_{MG}) is computed from the *disjunction* of ω_{SG_1} and ω_{SG_2}. The subjective opinion on MG represents the *belief* that an attacker can successfully achieve their main goal, the *disbelief* that an attacker can successfully achieve their main goal, and the *uncertainty* degree about the distribution of these belief and disbelief values.

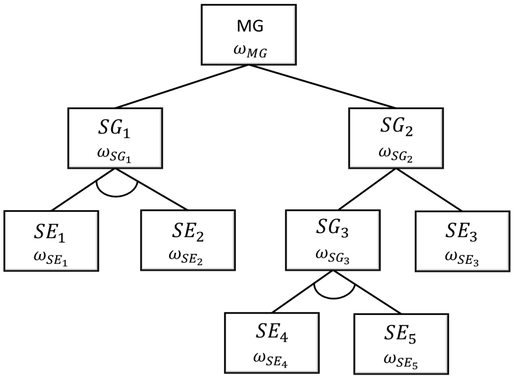


Fig. 2: A Subjective Attack Tree (SAT) model uses subjective opinions as input parameters to capture uncertainty degrees about the events' likelihoods. Here, ω_i is a subjective opinion capturing aspects of the likelihood of event i.

5 Security Events Evaluation using Subjective Logic

In this section, we propose an approach to derive subjective opinions about security events, using the evaluation and two criteria proposed in [1]. In our approach, uncertainty about likelihoods of security events (as discussed in the introduction) is due to uncertainty about the evaluation of the two criteria. We first need to consider quantitative values describing likelihood levels from combining technical difficulty levels with the vulnerability levels. An example of mapping qualitative scales into corresponding quantitative values is shown in Table 1.

Table 1: Corresponding quantitative values to likelihood qualitative scales.

Rating	Qualitative scales	Quantitative values	Description
1	very low	[0.1-0.2]	highly unlikely to occur
2	low	[0.2-0.4]	will most likely not occur
3	moderate	[0.4-0.6]	possible to occur
4	high	[0.6-0.8]	likely to occur
5	very high	[0.8-1.0]	highly likely to occur

5.1 The Two Criteria Evaluation

We mentioned in the introduction that it is often difficult to precisely determine the level of a vulnerability or technical difficulty of an attack. We propose a novel way to model uncertainty about the evaluation of these two criteria, allowing one to derive subjective opinions about security events, used then as input parameters in SATs.

Since each criterion specifies a number of categories (i.e., levels), where only one category represents the truth value in a given case, these categories thus represent the state space of a given criterion, and accordingly, the two criteria can be thought of as two frames of discernment. The state space of a vulnerability level is $VL = \{e, m, h\}$, and the state space of the technical difficulty is $TD = \{t, m, d, vd\}$. In our approach, security analysts need to assign values from the interval [0, 1] to each category, denoting their *degrees of belief* that each category represents the truth value. In addition, they complement these degrees by an uncertainty mass, provided that the sum of all the beliefs and uncertainty mass must equal to one. Furthermore, they assign a base rate to each category, as prior probability in absence of evidence, where the sum of the base rates must equal to one. Unless specified otherwise, we assume a uniform distribution for the base rates— the base rate of each category in the vulnerability level's frame of discernment is given as $1/3$ (≈ 0.33), and as $1/4$ ($= 0.25$) in the technical difficulty's frame of discernment. Fig 3 shows three examples of belief assignments in a vulnerability level's frame of discernment given different uncertainty masses about beliefs distribution.

b_e	b_m	b_h	u_{VL}
0	0	1	0

(a)

b_e	b_m	b_h	u_{VL}
0	0	0.8	0.2

(b)

b_e	b_m	b_h	u_{VL}
0	0	0	1

(c)

Fig. 3: Examples of belief assignments in a vulnerability level's frame of discernment: (a) the vulnerability level is 'high' with 100% confidence (i.e., 0 uncertainty), (b) the vulnerability level is 'high' with 0.2 uncertainty, and (c) complete uncertainty about the vulnerability level. u_{VL} stands for the uncertainty mass in the frame of discernment.

5.2 Evaluation Rules for Security Events

Because of uncertainty about the values of the two criteria, computing final opinions about security events directly using an evaluation matrix such as of Fig 1 is too complex. This requires to do multiplication of each likelihood level with the values of the corresponding combination of a Vulnerability level and technical difficulty, meaning that we need to perform twelve calculations. To facilitate the computation of subjective opinions, we propose a simple specification that translates the matrix in Fig 1 into a form of rules, calling them *evaluation rules*. The specification compacts the matrix information using a simple syntax such as

$$(VL = value_{VL}) \wedge (TD = value_{TD}) \Rightarrow^{W} SE, \tag{2}$$

where $VL = value_{VL}$ denotes the level of a vulnerability, $TD = value_{TD}$ denotes the technical difficulty of an attack, *SE* denotes security events for evaluation, and $\wedge$ is the conjunction symbol (i.e., AND). $VL = value_{VL}$ and AD $= value_{TD}$ are called the *antecedents* of the rule, while *SE* is the *consequent*. Further, the rule is given some form of weight, represented by W above the implication symbol $\Rightarrow$, denoting the likelihood level of *SE* occurrence given the values of the antecedents. The rule's weight corresponds to a cell value in a matrix. For example, the evaluation of a security event given that the vulnerability level is *easy E* and technical difficulty is *difficult D* according to the matrix in Fig 1 can be formulated as (assuming that the quantitative value corresponding to rating 3 is 0.5):

$$(VL = easy) \wedge (TD = difficult) \Rightarrow^{0.5} SE$$

When the same evaluation (i.e., the same unique likelihood level) is given for more than one combination, we use the *union* operator ($\cup$) as follows

$$\begin{aligned} &(VL = value_{VL} \wedge TD = value_{TD})_{comb_1} \\ &\cup (VL = value_{VL} \wedge TD = value_{TD})_{comb_2} \\ &\cup \cdots \cup (VL = value_{VL} \wedge TD = value_{TD})_{comb_n} \Rightarrow^{W} SE, \end{aligned} \tag{3}$$

where $comb_1$ denotes the first combination of vulnerability level and technical difficulty, $comb_2$ denotes the second combination, and so on, and $comb_1 \neq comb_2 \neq \ldots \neq comb_n$.

This rule can be further simplified. We may use the relation symbols of $\leq$ and $\geq$ to express a group of consecutive cells whose combinations are less than or equal (or greater than or equal) a certain level of vulnerability, technical difficulty, or both of them (with the assumption that there is a total order on the values of the two criteria). For example, the combinations of (hard H, trivial T) and (hard H, moderate M) in Fig 1 can be expressed as $(VL = hard) \wedge (TD \leq moderate)$. $TD \leq moderate$ in this example means that the technical difficulty's values are *moderate* and *trivial*. Accordingly, the evaluation rule is written as (with 0.3 corresponds to rating 2):

$$(VL \leq medium) \wedge (TD \leq moderate) \Rightarrow^{0.3} SE. \tag{4}$$

As in Eq 3, the union symbol $\cup$ can be also used to link antecedents that involve the relation symbols $\leq$ and $\geq$ in their expressions. For example, in Fig 1, since the rating 4 (0.7 in our quantitative example) is given for $(VL \leq medium \wedge TD = trivial)$ and for $(VL = easy \wedge AD = moderate)$, we formulate the evaluation's rule as

$$(VL \leq medium \wedge TD = trivial) \cup (VL = easy \wedge TD = moderate) \Rightarrow^{0.7} SE.$$

Based on the above discussion, we generalise Eq 2, Eq 3, and Eq 4 to obtain a more general form of security events evaluation as follows

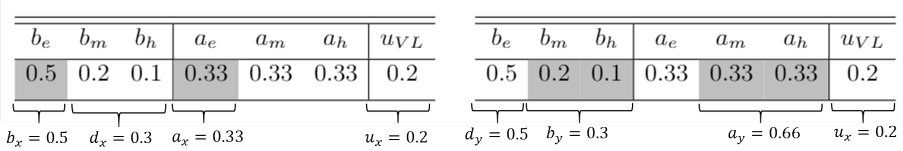
$$\begin{aligned} &(VL \odot value_{VL} \wedge TD \odot value_{TD})_{comb_1} \\ &\cup (VL \odot value_{VL} \wedge TD \odot value_{TD})_{comb_2} \\ &\cup \dots \cup (VL \odot value_{VL} \wedge TD \odot value_{TD})_{comb_n} \Rightarrow^{W} SE, \end{aligned} \tag{5}$$

where $\odot$ is any relation symbol from the set $\{=, \leq, \geq\}$, $comb_1$ denotes the first combination of likelihood level and technical difficulty, $comb_2$ denotes the second combination, and so on, and $comb_1 \neq comb_2 \neq \ldots \neq comb_n$.

5.3 Computing Final Opinions about Security Events

We use the proposed evaluation rules to derive subjective opinions about security events. We first need to evaluate each single antecedent in a rule (e.g., $VL = hard$ and $TD \geq difficult$) using the belief assignments in the frames of discernment of the two criteria . Next, we evaluate the combined antecedents in a rule (e.g., $(VL = easy \wedge TD \leq moderate) \cup (VL = medium \wedge TD = trivial)$) using the corresponding operators of $\wedge$ and $\cup$ in Subjective Logic. The symbol $(\wedge)$ is used to link two antecedents of different types to express a combination of technical difficulty and vulnerability level. The symbol $(\cup)$ is used to link multiple combinations of the same evaluation.

First, each single antecedent is evaluated by deriving a *binomial* opinion about it since their states can be either true or false. To derive a binomial opinion about an antecedent of the form $CT = value_{CT}$, where $CT \in \{VL, TD\}$ (i.e., the criterion type), and $value_{CT}$ is a category belongs to a given criterion, the



(a) $\omega_X = < 0.5, 0.3, 0.33, 0.2 >$

(b) $\omega_Y = < 0.3, 0.5, 0.66, 0.2 >$

Fig. 4: Deriving binomial opinions about two antecedents (a) $X : VL = easy$, and (b) $Y : VL \geq medium$.

belief mass of the binomial opinion takes exactly the same belief mass associated to the category $value_{CT}$ in the frame of discernment, and the disbelief mass of the binomial opinion is equal to the *sum* of all beliefs assigned to the other categories. The uncertainty of the subjective opinion takes the same uncertainty mass associated to the whole frame of discernment. Further, the base rate of the binomial opinion is exactly the same base rate associated to that category. Fig 4 shows an example beliefs and base rates assignments in a *VL*'s frame of discernment. Suppose we want to derive a subjective opinion about $VL = easy$, this process is demonstrated in Fig 4 (a).

To derive a binomial opinion about an antecedent of the form $CT \odot value_{CT}$, where $\odot \in \{\leq, \geq\}$, the belief mass of the binomial subjective opinion is the *sum* of all beliefs assigned to the categories starting from $value_{CT}$ and higher than this category in case of $\odot = \{\geq\}$, or the *sum* of all beliefs assigned to the categories starting from $value_{CT}$ and lower than this category in case of $\odot = \{\leq\}$. The disbelief mass of the binomial opinion takes the *sum* of all beliefs assigned to the remaining categories. The uncertainty of the binomial opinion takes exactly the same uncertainty mass associated to the whole frame of discernment. Further, the base rate of the binomial opinion is the *sum* of all base rates assigned to the categories starting from $value_{CT}$ and higher than this category in case of $\odot = \{\geq\}$, or the *sum* of all base rates assigned to the categories starting from $value_{CT}$ and lower than this category in case of $\odot = \{\leq\}$. Fig 4 (b) demonstrates the process of deriving a binomial opinion about $VL \geq medium$.

As a next step, we derive a binomial opinion about the antecedents. In Subjective Logic, the symbol $\wedge$ corresponds to the multiplication (conjunction) operator, and the symbol $\cup$ corresponds to the addition operator. Following this, we derive a final opinion about a security event. This is achieved by multiplying the obtained subjective opinion about the antecedents with the rule's weight. Because of uncertainty about the two criteria values, different evaluations (i.e., different subjective opinions) are obtained for security events, and the number of evaluations is equal to the number of rules.

Let r_i be an evaluation rule, where $1 \leq i \leq n$, and n is the number of evaluation rules, and the rule's strength is denoted by W_{r_i}. Let also SE be a security event for evaluation. According to Eq 5 and the operators of conjunction $(\cdot)$ and addition $(+)$, the subjective opinion on the security event SE is computed

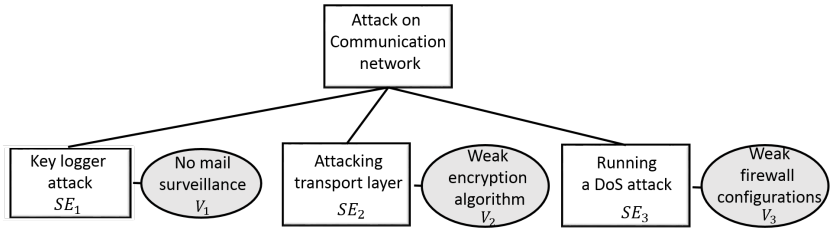

Attack on Communication network
Key logger attack SE_1
No mail surveillance V_1
Attacking transport layer SE_2
Weak encryption algorithm V_2
Running a DoS attack SE_3
Weak firewall configurations V_3

Fig. 5: The security actions and associated vulnerabilities (ovals) in Example 1.

from rule r_i as follows

$$\begin{aligned}\omega_{SE_i} =& ((\omega_{VL\odot value_{VL}} . \omega_{TD\odot value_{TD}})_{comb_1} \\ & + (\omega_{VL\odot value_{VL}} . \omega_{TD\odot value_{TD}})_{comb_2} \\ & + \cdots + (\omega_{VL\odot value_{VL}} . \omega_{TD\odot value_{TD}})_{comb_n}).W_{r_i}.\end{aligned} \tag{6}$$

To perform multiplication of a subjective opinion (about the antecedents) with a single value (the rule weight), we multiply each of the belief mass and base rate of the subjective opinion with the rule weight while maintaining the same uncertainty degree. This process ensures that the projected probability of the resulting subjective opinion (about a security event) is the same as if we multiply the projected probability of the subjective opinion about the antecedents with the rule weight. Formally, assuming $\omega_x = \langle b_x, d_x, u_x, a_x \rangle$ is the subjective opinion about antecedents in a rule of wight y ($W = y$), then a subjective opinion about a security event SE is computed as $\omega_x = \langle b_x.y, d_x, u_x, a_x.y \rangle$.

Finally, because of different possible outcomes obtained for a security event, we choose only one outcome to represent an input parameter in a SAT. In this paper, we work under the *most expected* risk scenario, by choosing the outcome that represents the most expected likelihood for a security event. For this purpose, we use the projected probability function (see Eq 1), which provides an estimate for the ground truth value of a variable by capturing the most likely value in presence of base rates.

Example 1. Suppose that in order to disrupt a communication network, the attacker needs to perform any of the following security actions: installing a key logger, attacking the transport layer, or running a DoS attack, via exploiting some existing vulnerabilities as shown in Fig 5. Suppose also the evaluation of security events is expressed by the following three rules:

$$r_1 : (VL \leq medium \land TD \leq moderate) \Rightarrow^{0.8} SE$$
$$r_2 : (VL = easy \land TD \geq difficult) \cup (VL = hard \land TD \leq moderate) \Rightarrow^{0.5} SE$$
$$r_3 : (VL \geq medium \land TD \geq difficult) \Rightarrow^{0.2} SE$$

Further, the beliefs assignments to each category in the frames of discernment of the level of each vulnerability and technical difficulty of each security event is given in Table 2. By deriving binomial opinions about the antecedents of the three rules, and using Eq 6 to compute subjective opinions about the security events, we obtain three possible subjective opinions for each security event (see Table 3). Having computed the projected probability of these subjective opinions to obtain the most expected value of each security event, we conclude that $\omega_{SE_1} = \langle 0.618, 0.252, 0.130, 0.264\rangle$, $\omega_{SE_2} = \langle 0.142, 0.569, 0.289, 0.660\rangle$, and $\omega_{SE_3} = \langle 0.567, 0.287, 0.146, 0.264\rangle$, and these would represent input parameters in Fig 5.

Table 2: Beliefs assignments in the frames of discernment of (a) the level of each vulnerability and (b) technical difficulty of each attack in Example 1.

(a)

Vulnerability	b_e	b_m	b_h	u_{VL}
V_1	0.15	0.60	0.05	0.20
V_1	0.00	0.15	0.70	0.15
V_2	0.30	0.50	0.10	0.10

(b)

Event	b_t	b_m	b_d	b_{vd}	u_{TD}
SE_1	0.85	0.05	0.05	0.00	0.05
SE_1	0.00	0.00	0.65	0.05	0.30
SE_2	0.20	0.60	0.05	0.00	0.15

Table 3: The possible subjective opinions about security events in Example 1.

Security event	Possible subjective opinions	Rule of derivation
SE_1	$\langle 0.618, 0.252, 0.130, 0.264\rangle$	r_1
	$\langle 0.047, 0.863, 0.090, 0.165\rangle$	r_2
	$\langle 0.009, 0.952, 0.039, 0.660\rangle$	r_3
SE_2	$\langle 0.009, 0.912, 0.079, 0.264\rangle$	r_1
	$\langle 0.053, 0.797, 0.150, 0.165\rangle$	r_2
	$\langle 0.142, 0.569, 0.289, 0.660\rangle$	r_3
SE_3	$\langle 0.567, 0.287, 0.146, 0.264\rangle$	r_1
	$\langle 0.068, 0.865, 0.067, 0.165\rangle$	r_2
	$\langle 0.011, 0.904, 0.085, 0.660\rangle$	r_3

6 Propagation of Subjective Opinions in SATs

So far, we have discussed the model of SAT and how to derive subjective opinions about security events as input parameters in the model. In this section, we discuss how these subjective opinions are propagated (through the gates of AND and OR) such that a subjective opinion on the root node can be then obtained.

Subjective opinions are propagated through AND gate using the *conjunction* operator. Let Z be an AND node in a SAT, with X and Y are its children. Let also $\omega_X = \langle b_x, d_x, u_x, a_x \rangle$ and $\omega_y = \langle b_y, d_y, u_y, a_y \rangle$ be the subjective opinions on X and Y, respectively. The subjective opinion on Z, ω_Z, is computed as $\omega_Z = \omega_x \cdot \omega_y$. Fig 6 (a) shows an example computation of a subjective opinion on event Z via AND gate.

Subjective opinions are propagated through OR gate using the *disjunction* operator. Let Z be an OR node in a SAT, with X and Y are its children. Let also $\omega_X = \langle b_x, d_x, u_x, a_x \rangle$ and $\omega_y = \langle b_y, d_y, u_y, a_y \rangle$ be the subjective opinions on X and Y, respectively. The subjective opinion on Z, ω_Z, is computed as $\omega_Z = \omega_x \sqcup \omega_y$. Fig 6 (b) shows an example computation of a subjective opinion on event Z via OR gate.

The operators of conjunction and disjunction on subjective opinions proved to be commutative and associative [9], and therefore the order of nodes (both AND and OR nodes) in an AT is not important.

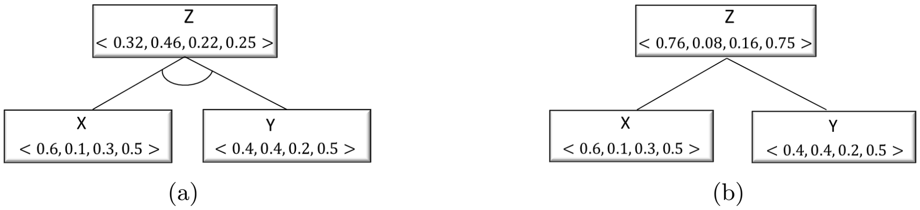


Fig. 6: Computing an opinion on event Z via (a) AND gate, and (b) OR gate.

7 Experimental Evaluation

We conduct an experimental evaluation to compare our approach with traditional probabilistic ATs, using the Stuxnet attack tree [1] as an illustrative example. To make the example simple, we consider only the operation of self-installation as demonstrated in Fig 7. Also, we omit the modelling of the vulnerability information about the security events, assuming their evaluations are obtained according to the two criteria and methodology we proposed in this paper, since the main goal of the section is to demonstrate why uncertainty should be taken into account when conducting risk analysis using models such as ATs.

We conduct three experiments, in each of which, we work with a different set of probabilities to compute the likelihood of the attack. We then start producing uncertainty about these probabilities. Uncertainty about a probability distribution is produced such that it affects a support to its belief mass only, a support to its disbelief mass only, or a support to both its belief and disbelief masses.

For a better study of the impact of uncertainty about the probabilities on the outcomes, we produce different degrees of uncertainty at each time of evaluation. We choose that, at each time, uncertainty about the probabilities is

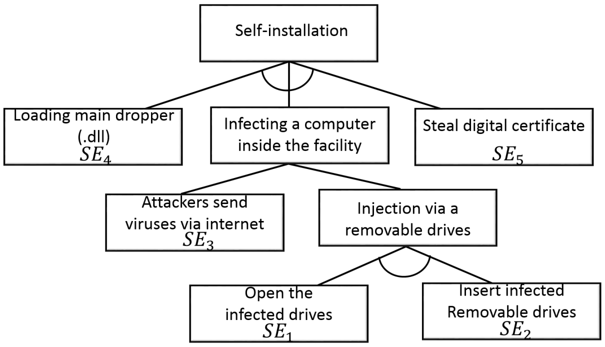

Self-installation
Loading main dropper (.dll) SE_4
Infecting a computer inside the facility
Steal digital certificate SE_5
Attackers send viruses via internet SE_3
Injection via a removable drives
Open the infected drives SE_1
Insert infected Removable drives SE_2

Fig. 7: Attack tree of "Stuxnet self-installation".

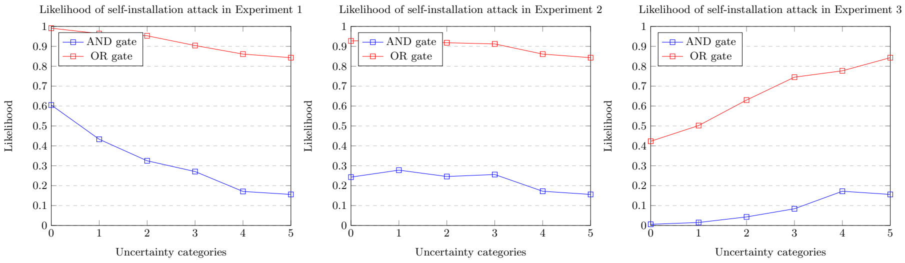
increased by at most %25, and for one time we consider the situation of complete uncertainty about the security events' probabilities. Here, we consider the following uncertainty categories: (1) $u_X \in [0.01, 0.25]$, (2) $u_X \in [0.26, 0.50]$, (3) $u_X \in [0.51, 0.75]$, (4) $u_X \in [0.76, 1.0]$, and (5) $u_X = 1.0$, where X is any security event in the given AT. Due to space limitation in this paper, we show only the set of probabilities and subjective opinions used in Experiment 1 (see Table 4). The set of probabilities used in the other two experiments are as follows: 0.3, 0.8, 0.6, 0.7, and 0.5 (for experiment 2), and 0.6, 0.9, 0.6, 0.1, and 0.1 (for experiment 3) for the security events SE_1, SE_2, SE_3, SE_4, and SE_5 in order. Uncertainty about these probabilities is produced in the same way as in Experiment 1.

Table 4: Probabilities and subjective opinions used in Experiment 1.

Uncertainty	SE_1	SE_2	SE_3	SE_4	SE_5
$u_x = 0$	0.7	0.86	0.6	0.8	0.9
$u_x \in [0.01, 0.25]$	$\langle 0.60, 0.25, 0.15 \rangle$	$\langle 0.65, 0.10, 0.25 \rangle$	$\langle 0.40, 0.40, 0.20 \rangle$	$\langle 0.65, 0.25, 0.10 \rangle$	$\langle 0.70, 0.08, 0.22 \rangle$
$u_x \in [0.26, 0.50]$	$\langle 0.50, 0.20, 0.30 \rangle$	$\langle 0.55, 0.10, 0.35 \rangle$	$\langle 0.30, 0.25, 0.45 \rangle$	$\langle 0.30, 0.20, 0.50 \rangle$	$\langle 0.60, 0.00, 0.40 \rangle$
$u_x \in [0.51, 0.75]$	$\langle 0.25, 0.05, 0.70 \rangle$	$\langle 0.35, 0.10, 0.55 \rangle$	$\langle 0.17, 0.20, 0.63 \rangle$	$\langle 0.15, 0.10, 0.75 \rangle$	$\langle 0.34, 0.00, 0.66 \rangle$
$u_x \in [0.76, 1.00]$	$\langle 0.20, 0.00, 0.80 \rangle$	$\langle 0.10, 0.05, 0.85 \rangle$	$\langle 0.00, 0.00, 0.10 \rangle$	$\langle 0.05, 0.00, 0.95 \rangle$	$\langle 0.01, 0.00, 0.99 \rangle$
$u_x = 1$	$\langle 0.00, 0.00, 0.10 \rangle$	$\langle 0.00, 0.00, 0.10 \rangle$	$\langle 0.00, 0.00, 0.10 \rangle$	$\langle 0.00, 0.00, 0.10 \rangle$	$\langle 0.00, 0.00, 0.10 \rangle$

In addition to the given AT structure of self-installation, we repeat the same above experiments for a modified structure in which AND gates are replaced with OR gates, and vice versa, the OR gates are replaced with AND gates (although this doesn't offer a real representation of the self-installation scenario, but we do so for demonstration purposes only, and therefore should not be taken as a real representation of the attack). We do swap between the gates in order to also study the outcomes in case that the target node of evaluation is of type OR.

Using the prorogation method of probabilities (discussed in literature) and propagation method of subjective opinions (discussed in this paper), we obtained probabilities for the self-installation attack and subjective opinions on it. To com-



(a) (b) (c)

Fig. 8: Likelihood of self-installation attack in the three experiments using AT (denoted by uncertainty category with label 0), and SAT models (denoted by uncertainty categories with labels 1 to 5 as defined in text).

pare the outcomes (likelihoods) from using both approaches, we assumed here the most expected scenarios when dealing with subjective opinions by computing their projected probabilities. Fig 8 shows the likelihoods of self-installation attack in each experiment when there is no uncertainty about the probabilities (AT model) and when there is uncertainty about them (SAT model) based on the five defined uncertainty categories (numbered from 1 to 5), and with different gate type of the root node.

In Experiment 1, and in case of AND gate of the root node, the likelihood of self-installation attack decreases as uncertainty about the probabilities increases, and the decrease is to somewhat sharp in case of total uncertainty about the probabilities, resulting in a reduction from 0.605 to 0.15625 (i.e., the difference in probability is approximately 0.448). Unlike the case of AND gate, the projected probabilities of the subjective opinions given that the root node is of type OR decease very slightly as uncertainty increases, and the difference in the probability when there is no uncertainty and when there is total uncertainty about the probabilities is only 0.147. Here, the effect of uncertainty about probabilities in this particular case is very small. Graph (a) of Fig 8 demonstrates that taking uncertainty about the probabilities into account when the root node is of type AND leads to very different results than in case of OR gate.

In Experiment 2, whether the root node is of type AND or OR, the results are not considerably different in case of AT or SAT model. The maximum difference in probability using both structures when there is no uncertainty about the probabilities and when working with total uncertainty about them is only 0.084.

In Experiment 3, both structures result in an increase in the likelihood of the attack as uncertainty increases. However, the increase is very high in case of OR gate, nearly 0.42 as probability difference when using the probabilistic approach and when $u_X = 1.0$, while it is slight in case of AND gate (only 0.159). The analysis here is opposite to the one in Experiment 1, where both gates lead to a decrease in the likelihood and such a decrease is sharper in case of AND gate than of OR gate.

Importantly, there are cases such that in the AT approach, the decision is to not protect the system, while it is the reverse in the SAT model. As an example with OR structure in Experiment 3, the security manager would only consider a protection mechanism against the attack if the probability is greater than 0.5. This example, in particular, and the results from Experiment 1, in general, clearly demonstrate the importance of modelling uncertainty about probabilities when conducting security risk analysis — *doing so can lead to completely different security decisions being made.*

8 Conclusions and Future Work

We developed a new model of attack trees, called a subjective attack tree, that takes second-order uncertainty about input parameters into account, via subjective opinions. We proposed an approach to derive subjective opinions security events based on two criteria, a vulnerability level and technical difficulty of an attack. Our approach involved development of evaluation rules using subjective logic. Propagation of subjective opinions has been also discussed. Finally, we evaluated our approach against traditional ATs, showing that SATs lead to different outcomes in contrast to ATs, leading to different decisions being made.

As future work, we will consider other criteria to evaluate likelihoods of security events, such as connectivity of systems, technology and communication protocols used, users' behaviour, etc. Further, the current work has presented the foundation of SATs with only one input parameter, i.e., likelihood. For effective risk and decision analysis, we will need to extend the model by incorporating countermeasures, allowing for additional parameters to be included, such as cost of attack, cost of countermeasure, impact, and so on. We will discuss the impact of uncertainty in selecting the optimal set of countermeasures, comparing the results with existing approaches, e.g., [6, 11, 18, 20].

References

1. Abdo, H.: Dealing with uncertainty in risk analysis: combining safety and security. Ph.D. thesis (2017)
2. Buldas, A., Gadyatskaya, O., Lenin, A., Mauw, S., Trujillo-Rasua, R.: Attribute evaluation on attack trees with incomplete information. Computers & Security **88**, 101630 (2020)
3. Buldas, A., Laud, P., Priisalu, J., Saarepera, M., Willemson, J.: Rational choice of security measures via multi-parameter attack trees. In: International Workshop on Critical Information Infrastructures Security. pp. 235–248. Springer (2006)
4. Buoni, A., Fedrizzi, M., Mezei, J.: A delphi-based approach to fraud detection using attack trees and fuzzy numbers. In: Proceeding of the IASK International Conferences. pp. 21–28 (2010)
5. Edge, K., Raines, R., Grimaila, M., Baldwin, R., Bennington, R., Reuter, C.: The use of attack and protection trees to analyze security for an online banking system. In: 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07). pp. 144b–144b. IEEE (2007)

6. Edge, K.S., Dalton, G.C., Raines, R.A., Mills, R.F.: Using attack and protection trees to analyze threats and defenses to homeland security. In: MILCOM 2006-2006 IEEE Military Communications conference. pp. 1–7. IEEE (2006)
7. Gordon, J., Shortliffe, E.H.: The dempster-shafer theory of evidence. Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project **3**, 832–838 (1984)
8. Gribaudo, M., Iacono, M., Marrone, S.: Exploiting bayesian networks for the analysis of combined attack trees (2015)
9. Jøsang, A.: Subjective logic. Springer (2016)
10. Jürgenson, A., Willemson, J.: Processing multi-parameter attacktrees with estimated parameter values. In: International Workshop on Security. pp. 308–319. Springer (2007)
11. Kordy, B., Mauw, S., Radomirović, S., Schweitzer, P.: Foundations of attack–defense trees. In: International Workshop on Formal Aspects in Security and Trust. pp. 80–95. Springer (2010)
12. Kumar, R., Stoelinga, M.: Quantitative security and safety analysis with attack-fault trees. In: 2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE). pp. 25–32. IEEE (2017)
13. Mauw, S., Oostdijk, M.: Foundations of attack trees. In: International Conference on Information Security and Cryptology. pp. 186–198. Springer (2005)
14. Moore, A.P., Ellison, R.J., Linger, R.C.: Attack modeling for information security and survivability. Tech. rep., Carnegie-Mellon Univ Pittsburgh Pa Software Engineering Inst (2001)
15. Opel, A.: Design and implementation of a support tool for attack trees. Internship Thesis, Otto-von-Guericke University Magdeburg (March 2005) (2005)
16. Pieters, W., Davarynejad, M.: Calculating adversarial risk from attack trees: Control strength and probabilistic attackers. In: Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance, pp. 201–215. Springer (2014)
17. Roy, A., Kim, D.S., Trivedi, K.S.: Cyber security analysis using attack countermeasure trees. In: Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research. pp. 1–4 (2010)
18. Roy, A., Kim, D.S., Trivedi, K.S.: Attack countermeasure trees (act): towards unifying the constructs of attack and defense trees. Security and Communication Networks **5**(8), 929–943 (2012)
19. Schneier, B.: Attack trees. Dr. Dobb's journal **24**(12), 21–29 (1999)
20. Wang, P., Lin, W.H., Kuo, P.T., Lin, H.T., Wang, T.C.: Threat risk analysis for cloud security based on attack-defense trees. In: 2012 8th International Conference on Computing Technology and Information Management (NCM and ICNIT). vol. 1, pp. 106–111. IEEE (2012)