TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Information Technology
Department of Informatics

IDK70LT

Maria Hents 111671IAPMM

# COST REDUCTION IN ATTACK TREE BASED QUANTITATIVE SECURITY ASSESSMENT

Master's Thesis

Supervisor: Aleksandr Lenin
Ph.D
Researcher

Tallinn 2016

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Informaatika instituut

IDK70LT

Maria Hents 111671IAPMM

# MAKSUMUSTE REDUTSEERIMINE RÜNDEPUUDELE PÕHINEVAS KVANTITATIIVSES RISKIANALÜÜSIS

Magistritöö

| | |
|---|---|
| Juhendaja: | Aleksandr Lenin |
| | Ph.D |
| | Teadur |

Tallinn 2016

# Declaration of Authorship

Hereby I declare that this master thesis, my original investigation and achievement, submitted for the master degree at Tallinn University of Technology has not been submitted before for any degree or examination.

__________________________ __________________________

Date Author's signature

# Abstract

As the provision of security is an expensive investment, it is important to have a clear vision of which protection measures should be introduced in order to make the system stay protected from harmful actions against it. Improved Failure-Free model uses an effective propagation method with cost reduction technique in order to calculate an expected attacker expense. By using the propagation method it is possible to calculate the cost of the primary threat. If there are common attacks in the attack scenario, the propagation method will not provide the exact result in many cases. In order to get the exact result, the cost reduction method is used, where dependent attacks are replaced by independent copies with lower cost. The main aim of this research is to check if cost reduction method will provide an exact result if the attack tree contains 2 or more common attacks. For conducted research, hypothesis have been formulated which will be proved or disapproved with the help of theorems. The accuracy of cost reduction method has been assessed in the process of research, conditions under which reduction defect exists have been found and the size of this defect has been calculated too.

This thesis is written in English and is 53 pages long, including 5 chapters and 21 figures.

# Annotatsioon

**Maksumuste redutseerimine ründepuudele põhinevas kvantitatiivses riskianalüüsis**

Turvalisuse tagamine on kallis investeering. Süsteemi kuritahtlike tegevuste eest kaitsmiseks on hädavajalik, et oleks selge ettekujutus sellest, milliseid turvameetmeid kasutada. Improved Failure-Free mudel kasutab efektiivset propageerimise meetodit koos maksumuse vähendamise tehnikaga. Seda kasutatakse, et arvutada ründajate eeldatavate kulusid. Propageerimise meetodit kasutades on võimalik arvutada primaarse ohu maksumus. Kui ründe stsenaariumis esinevad sõltuvad ründed, siis propageerimise meetod ei anna täpset tulemust. Täpse tulemuse saamiseks kasutatakse maksumuse vähendamist, kus sõltuvad ründed asendatakse väiksema maksumusega sõltumatute koopiatega. Selle töö põhiülesanne on kontrollida, kas maksumuse vähendamise meetod annab täpse tulemuse kahe või rohkem sõltuva ründe korral. Uurimistöö käigus püstitati hüpoteesid, mis võiks olla tõestatud või ümberlükatud teoreemide abil. Uurimistöös anti täpsuse

hinnang maksumuse vähendamise meetodi jaoks, tuletati tingimused, mille korral viga esineb arvutustes, ja arvutati selle vea suurus.

Lõputöö on kirjutatud inglise keeles ning sisaldab 53 lehekülge teksti, 5 peatükki, 21 joonist.

# Abbreviations and Symbols

| Abbreviation | Definition |
|---|---|
| WMSAT | Weighted Monotone Satisfiability |
| PSPACE | The set of decision problems that can be solved by a Turing machine using a polynomial amount of space |

| Symbol | Definition |
|---|---|
| $\phi$ | Boolean formula |
| $F$ | Boolean formula |
| $X$ | The set of variables of $F$ |
| $x_i$ | A variable in $X$ |
| $FG$ | Conjunction $F \wedge G$ |
| $F + G$ | Disjunction $F \vee G$ |
| $\partial F$ | Boolean Derivative |
| $\mu$ | min-tem of $F$ |
| $w(\phi)$ | Weight function - weight of $\phi$ |
| $\phi'$ | Function $\phi$ after cost reduction |
| $x'$ | An independent variable with reduced cost |
| $NP$ | Nondeterministic polynomial time |

# Contents

# List of Figures

# 1 Introduction

Information technologies are developing with a fast pace. Nowadays it is often necessary to think of qualified provision of security for systems. As the provision of security is an expensive investment, it is important to have a clear vision of which protection measures should be introduced in order to make the system stay protected from harmful actions against it. A threat against an enterprise can be structurally represented by a structure called ”attack tree”. With the help of structured description of threat that consists of several elementary steps, attack trees’ structure allows to conduct an assessment of interrelations between separate components. In this structure a root-node of the tree is a primary threat and the tree leaves are elementary actions called attacks. Improved Failure-Free model uses an effective propagation method with cost reduction technique in order to calculate an expected attacker expense. By using the propagation method it is possible to calculate the cost of the primary threat. If there are common attacks in the attack scenario, the propagation method will not provide exact result in many cases. In order to get the exact result, the cost reduction method is used, where dependent attacks are replaced by independent copies with lower cost. The main aim of this research is to verify if the cost reduction method will provide an exact result if the attack tree contains 2 or more common attacks.

In the process of research it has been discovered that if the attack tree contains 2 or more common attacks the propagation method does not give an exact result. A counter-example was found which then served as the basis for calculating the result of propagation method after applying the cost reduction technique. The defect which occurs during the result calculation for the attack tree with common attacks can lead to additional investments, as the attacker expenses can be understated. This thesis is focuses on possible ways how to make calculations more accurate.

Before using the propagation method, it is necessary to replace all dependent attacks by independent ones. The cost reduction technique is applied for getting rid of common attacks, where dependent attacks are replaced by several independent copies with lower cost. The process of getting rid of common attacks is an iterative process, where in every subsequent iteration, attack tree contains less common variables. During the research, the hypothesis ”given a Boolean function $F$ with 2 or more common variables, there always exists an optimal distribution of weights of common variables, that would give precise result", has been disapproved. Thus, we conclude that in the general case, reduction defects are present in calculations. Conditions, under which the reduction defect is present, have been discovered. It turned out that reduction defect is bounded, which allows us to perform meaningful marginal analysis. Upper and lower bounds of the reduction defect have been found in this thesis.

## Outline of the thesis

The structure of the thesis is following:

**Chapter 1** provides an overall impression about the contents of the thesis and also shortly describes its aims, methods of their achievements and received results.

**Chapter 2** provides theoretical information of security modelling. Chapter 2 tells about threat description, security definition, adversarial model and methods of security assessment.

**Chapter 3** is describes attack tree based quantitative security assessment methods, related to the present research.

**Chapter 4** outlines the contribution of the author. The hypothesis that there always is such optimal weight's distribution for the function with 2 or more common variables such that the cost reduction method will provide the exact result has been disapproved. Apart from that conditions under which the calculation defect exists, have been discovered. Two theorems proving that reduction defect is bounded, have been created.

**Chapter 5** sums up and describes the research results, outlines of unsolved questions, and provides directions for future research.

# 2 Theoretical Background

## 2.1 Security Modeling

Scientific studies can be conducted in a variety of ways. If there is any physical object, it means that experiments could be made, necessary measurements could be taken and parameters could be calculated for it. A physical object can be studied if it has the following qualities:

- observability (object should be observed)
- controllability (object should be controllable during different experiments)
- measurability (object's properties should be measurable in order to provide meaningful result)

During the studies of certain objects, some difficulties may arise as a result of the fact that one or more properties can not meet the conditions. For instance, weather, as many other natural phenomena, is an object which we observe, measure, but can not control. There is no need in conducting physical experiments if they have extremely high cost. We will not destroy satellites just to find out how resistable they are to crashes.

Security is an object being studied in this research. In this case, the object ”security” does not meet such necessary conditions as observability and measurability. We can not measure security, as such instruments do not exist. The solution would be modeling the object and calculate properties of this particular object. The model is a reflection of the real object. By creating the model, only parts of the object, relevant for the research, are modelled.

In one of his publications, Karl Popper[1] claims that every modeling should have an opportunity to show that an implemented method for calculation or the model is not correct. Such an approach is called ”falsification”. It is possible to falsify the result of security assessment only when a real attack occurs. It is possible only to guess that the result is correct before the attack takes place. If it is assumed that the results of calculation show that the system is secure, but the attacker managed to attack it, the conclusion can be made that the methods of calculations were not correct. Another example of falsification is an attack which has been successfully conducted by a method which has not been described here in this model. Therefore the model of the object of research was modelled incorrectly.

In his work, Lenin A.[2] decribed the structure of such a model. The typical model consists of four parts:

- Description of threats
- The model of an adversary
- Definition of security
- Methods of security assessment

Computational methods verify if the model is secure w.r.t. the provided definition of security in the model.

### 2.1.1 Description of threats

Rational profit-oriented attacks are the most concerning for enterprises. The attack will be considered as a profitable one for the attacker in the case when expenses for committing the attack will be less than potential profit. If expenses for conducting the attack will exceed the profit it means that this attack does not make any sense for rational attackers.

The structure of the attack introduced in the form of an ”attack tree” gives an opportunity to conduct the quantitative analysis of security. The attacker, by committing an attack, is executing a chain of actions corresponding to a certain strategy. The attack strategy is formed by such factors as motivation and availability of necessary resources. The root node of the tree describes the primary threat, which can cause harm to organization. The cost of the primary threat for the attacker is the cost of preparation and launching the attack scenario. An example of the attack tree is shown in Figure 2.1.

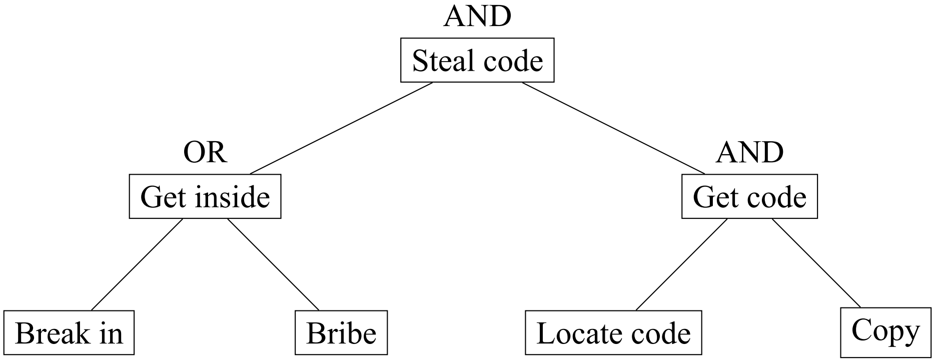

Figure 2.1: Attack tree nodes

In the example above, the attacker can steal the code in a variety of ways.

**Case 1.** Break in AND Locate code AND Copy
**Case 2.** Bribe AND Locate code AND Copy

### 2.1.2 The model of an adversary

There is a great number of various threats in the present world. For instance, we can not completely protect ourselves from the harm of natural disasters. Whereas human made harm is quite more predictable. We can model the human behavior and choose methods of protection which could influence the attacker. Rational attackers that get profit and thoroughly plan the attack are the most dangerous ones in industry[3]. Behavior of a rational attacker is shown in Figure2.2.

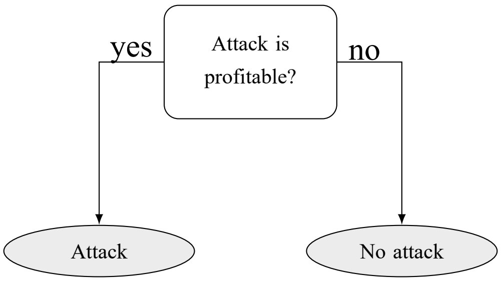

Figure 2.2: Behavioral model of a rational attacker

### 2.1.3 Definition of security

In terms of information security, the factor ”security” can be seen as a condition where all the risks are minimal. Risk is a definition widely used in various spheres of activity. Dan Ionita[4] has outlined 5 various types of risk.

**Class 1.**

$$Risk[Threat, Asset] = Likelihood[Threat] \times Vulnerability[Threat, Asset] \times$$

$$\times Impact[Threat, Asset]$$

This is a classic formula of calculating risk, where probability of threat occurrence is considered as well as possibilities of causing damage by this threat. This is commonly used in most general-purpose risk assessments.

**Class 2.**

$$Risk[Threat, Asset, Requirements) = Vulnerability(Threat, Asset] \times$$

$$\times Impact[Threat, Requirements]$$

By considering the consequences of the threat and possibility of threat's uprising we can define the correspondence of the object to previously introduced standards of security. (like ISO/IEC 27002).

**Class 3.**

$$Risk[Threat, Asset] = Likelihood[Threat, Asset] \times AverageLoss[Threat, Asset] =$$

$$= AnnualLossExpectancy[Threat, Asset]$$

This class of risk is used for defining risks in financial area. For calculating the risk in this particular case we consider the possibility of causing damage and loss connected with this threat. The risk is calculated by considering certain period of time – for instance one year.

**Class 4.**

$$Risk[Threat, CriticalAsset] = Vulnerability[CriticalAsset] \times$$

$$\times Impact[Threat, CriticalAsset]$$

This approach is used for providing Security-Critical systems. Such risk is calculated by combining all parameters being under attack and considering the possibility of causing damage to each of them. For instance, this approach is implemented for medical systems or in aviation-space technology.

**Class 5.**

$$Risk[Incident, Asset] = Likelihood[Incident] \times Impact[Incident, Asset]$$

This approach is based on traditional interpretation of risk for further security analysis. In this case an average frequency of threat uprising and consequences is considered. This approach is used for calculating risk that the object can face in connection with crash of on-board computers caused by ecological factors. Unlike Class 1, the risk is calculated without the presence of possibility of threat's uprising, so that it is limiting the area of implementing this method.

In research of Lenin A.[2] the main harm for organization is formed by two components:

- risk
- security investment

Research that represents the basis for this thesis use the following formulas for calculating the risk and damage:

$$Risk[Threat] = ProbabilityOfSuccess[Threat] \times Loss[Threat]$$

$$Loss[Threat] = Risk[Threat] + SecurityInvestment[Threat]$$

where $Threat$ marks potential threat, $ProbabilityOfSuccess[Threat]$ - probability of success, $Loss[Threat]$ - loss caused by the threat and $SecurityInvestment[Threat]$ - investments in security.

For analyzing the security of the system the model should be created, risks should be measured and necessary conclusions should be made. In case when the risk is high, additional necessary security measures must be implemented and the risk must be recalculated taking into consideration new methods of protection. Risk calculation should be conducted until the moment when the result meets certain criteria. By decreasing the risk, the so-called ”residual risk” still remains, which may be covered by insurance, or accepted.

### 2.1.4 Methods of security assessment

Quantitative security assessment has been studied by many researchers. Security is such object which cannot be measured, but it can be calculated in numerical way with the help of various methods[5, 6]. By taking the result of analysis into account, we can find out whether the studied model of the system is secure in accordance with definition of security in the model. It has been mentioned before that providing security is an expensive investment, therefore several reliable methods are necessary for calculating existing or predicted level of security. Reliable methods should not provide false-positive results. Efficiency of the chosen method can depend on such factors as the time for completing the calculation, the accuracy of the result, the speed of calculations. While calculating the exact result, possibility of mistake is considered. Obtaining false-positive results is less secure result than the false-negative ones. It is connected with the fact that having obtained a false-negative result additional measures of security are implemented. However, from the management point of view, false-negative results lead to additional investments. Methods of calculating exact results are complicated and tend to belong to the PSPACE. In practice it is enough to get approximate results. Such methods of calculation are reasonably faster.

The terminology of ”attack trees” introduced by Shneier[7] is widely used these days. He offered to describe threats in a structured manner and represent them as trees having certain structure where the primary threat is the root-node. Attack tree based security analysis can be classified into single-parameter and multi-parameter analysis. In the case with a single parameter, we can, for instance, assess possibility or impossibility of an attack scenario. Therefore, for instance, each leaf is given a status of possibility or impossibility of the action. The meaning of the root AND node will be ”possible” only if the action of each child-nodes is possible. In other words, if at least one of conditions is not completed, the aim will not be achieved. The OR node gives several alternative opportunities for achieving the aim. The OR node gets a ”possible” status when at least one of the child-nodes is a possible action. If all child-nodes of the OR node are labelled ”impossible”, the node is consequently also ”impossible”.

Apart from possibility or impossibility of actions, there are other parameters. Some of the parameters are the following:

- possible/impossible
- easy/difficult
- legal/illegal
- expensive/inexpensive
- fast/long
- special equipment required/no special equipment

The most significant factor for analysis of assumed rational attackers is the cost of planned attack. The overall cost of attack scenario is calculated in the root node of the tree by calculating all logical actions leading to the achievement of attacker goal. Therefore, for the OR node, the attack with minimal cost is chosen by the rational attackers. The cost of an AND node is the sum of the costs of the sub-attacks. Figure 2.3 shows the example of an attack tree where the values for cost each attack is marked.

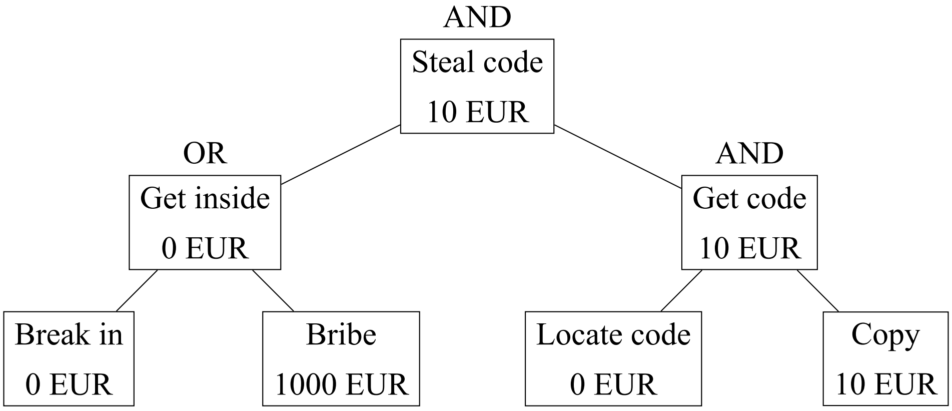

Figure 2.3: Cost of attack

The minimal cost of the attack scenario shown in Figure 2.3 is around 10 €.

Quantitative metrics, used for security analysis can be combined. By combining the parameters we can calculate expected profit of the attacker. By considering only ”expenses/cost” parameter, we could calculate the minimum cost of the whole scenario. By considering ”success probability” parameter, we could calculate the success probability of an attack scenario.

# 3 Related Work

Attack trees have been used for several decades. At the beginning of the 1980-s such structure was used for analyzing the risk of systems' fault and was called fault trees[8]. New forms of analysis were being invented. Therefore, a new structure of analysis called ”threat logic trees”[9] appeared in 1991.

Attack trees have become the foundation for various researches. Several amount of models have been created during last ten years which help to develop information protection. The publication Mauw et al.[10] published in 2005 has become the foundation of attack trees development. In their article, the authors suggested to abstract from the conception of structural attack description which had been suggested by Schneier[7] earlier and to study the attack with the set of possible attacks, ignoring grouping and connections between the components. Such sets got the name of ”attack suite”. By using the chronological order the models related to current research can be described in the following way:

- Multi-parameter Model (2006)
- Parallel model and Serial model (2008 - 2010)
- Fully-Adaptive Model and Failure-Free Model(2012)
- Improved Failure Free Model (2013)
- Improved Failure Free Model with Limited budget (2014)

The structure of the related models is described in the following sections.

## 3.1 Multi-parameter Model

Multi-parameter Model[3], described by Buldas et al. in 2006, is based on the assumption that the rational attacker always makes attempts to maximize the average payoff. The basis of the model is an economical approach. The main task of the methods of the risks' analysis is to define how well the organization is protected from attacks. Protection methods from unwanted interference is an expensive investment. In order to evade excessive investments into security, it should be carefully analyzed how to choose the most effective ways of protection. The so-called gain-oriented attacks have a clear quantitative estimation in most cases. Exceptional cases are the ones where the attacker is not using certain template which makes the process of applying methods of protection more difficult[11]. The multi-parameter model, like in all further models, assumes rational attackers that would like to get profit. Protection of enterprises is firstly directed against rational attacks where the risk has its maximum level. The authors state that in order to assess security, it is sufficient to assess adversarial utility. If the adversarial utility is a negative or a zero value, it means possible attacks are not likely to bring any profit. If adversarial utility has a positive value, it means that the attacker has a motivation to attack, as the attack will be profitable. This analysis considers profit of the attacker as well as possible consequences. Consequences of the attacking the system (as a rule) are the discovery of the attacker and the punishment in accordance with the crime committed. The drawback of the Multi-parameter Model is that the model cannot be used in the case when actions depend on each other. In the same way, another drawback is the limit of actions of the attacker. The model assumes completion of all actions in accordance with a previously created plan. According to the model, actions are completed step-by-step, not considering results received at each step. In the real life the attacker can act in the same way in accordance with the plan, but the sequence of actions may change, as each next step is completed on the basis of results received during the previous steps.

## 3.2 Parallel Model and Serial Model

The Parallel model [12] was created by Jürgenson et al. in 2008. Main principles of Mauw et. al.[10] were the basis of the model, as it works with the Boolean function of the attack tree and does not depend on the structural representation of the attack tree. The model provides a possibility to assess the exact expected result of the attacker using multi-parameter attack tree. According to this model, an attacker calculates all possible ways of attacking by looking at various combinations called attack suites. Having assessed all the results, he chooses combination which has given the most profitable result and at the same time launches all attacks from the attack suite. Similarly to the Multi-parameter Model[3], parallel model still limits actions of the attacker.

Serial model [13] is a continuation of parallel model. Process of making decisions by the attacker is the basis of it. The attacker completes the attack steps one by one. Taking previously achieved results into consideration, the attacker can make decision to skip one elementary attack or to finish completing the attack. Despite Serial model gives more freedom of actions to the attacker, the model still limits adversarial actions and does not give a possibility to choose the sequence of actions.

## 3.3 Fully-Adaptive Model and Failure-Free Model

Fully-Adaptive Model and Failure-Free Model[14] were created by Buldas et al. in 2012. Unlike previous models, the Fully-Adaptive Model allows the attackers to act in a fully adaptive way by launching elementary steps of the attack in any order and takes the results of the previous trials into consideration. However, the attacker can launch an attack step not more than once. This approach is not close enough to real situations. Failure-Free Model means that the completion of attack steps is used until successful result is achieved. Failure-Free Model is a model similar to the Fully-Adaptive Model. The only difference is that in the Faiure-Free Model the success probability equals 1. The Fully-Adaptive Model and the Failure-Free Model, as well as the Multi-parameter Model[3], are based on the analysis of the upper bound of adversarial utility.

## 3.4 Improved Failure-Free Model

In his publications, Buldas et al. described the fully-adaptive model[14], the main principle of which is completion of attacks in random order considering previous the results of the trials as the main basis. Unlike fully-adaptive model, the Improved Failure-Free Model[15, 2] is a model that allows the opponent to complete attacks even if the original approach led to failure and to continue the attack, even if the attack was discovered. Model which excludes possible ”unsuccess” turned out to be more simple for analysis in comparison with the fully-adaptive model. Therefore, if there is a need to complete the sequence of actions, each of which should definitely be completed, there is no difference in order.

The Improved Failure-Free Model, as well as the fully-adaptive model, consists of 4 parts, as described in Part 2.1. The threat model in the Improved Failure Free Model is an attack tree which represents a primary threat. The model considers rational attackers only. Optimal strategy of the attack will be the cheapest one among other possible strategies. Security of the system, in the context of the studied model, is achieved by decreasing attractiveness of the system for rational attackers. The system can be called secure when required investments for the attack completion go beyond potential profit.

There are several methods of calculation used by the Improved Failure-Free Model. These methods are the exact method and the propagation method. The exact method gives an accurate result, but the calculation process has an exponential complexity in the worst case. By using the exact method the accurate result is calculated using the following formula:

$$w(\phi) = \min\{w(x_i) + w(\phi \mid_{x_i=1}), w(\phi \mid_{x_i=0})\},$$

where $x_i$ is an elementary attack step, $w$ is the cost function and $\phi \mid_{x_i}$ is the Boolean

”derivative” of $\phi$.

The propagation method is more effective for calculation. That calculate upper bounds of adversarial utility. Utility upper bounds are calculated using the following rules:

$$w(x_1 \vee ... \vee x_n) \leq \min\{w(x_1), ..., w(x_n)\}, \tag{3.1}$$

$$w(x_1 \wedge ... \wedge x_n) \leq w(x_1) + ... + w(x_n) \tag{3.2}$$

Therefore, in the case of a conjunction, the weights of all sub-attacks are summed up and in the case of a disjunction the weights of the cheapest sub-attack is chosen.

The propagation method propagates adversarial expenses in a bottom-up manner using the formula (3.1) and (3.2). The overall cost of the attack is calculated in the root node and represents expenses $\varepsilon$, necessary for completing the attack. In order to calculate upper bounds of adversarial utility, it is necessary to calculate the lower bound of expenses. Utility is calculated using the formula $U = (P - \varepsilon)$, where $U$ is adversarial utility and $P$ is adversarial profit.

When applying the propagation method, the result will be exact for independent attack trees, which do not contain common attacks. In the case of independent attack trees, the following equations are used:

$$w(x_1 \vee ... \vee x_n) = \min\{w(x_1), ..., w(x_n)\},$$

$$w(x_1 \wedge ... \wedge x_n) = w(x_1) + ... + w(x_n)$$

Figure 3.1 demonstrates a dependent attack tree with one common attack, where the propagation method does not provide exact result. An optimal solution, that is also the lower bound of expenses, in this case will be the cheapest solution.

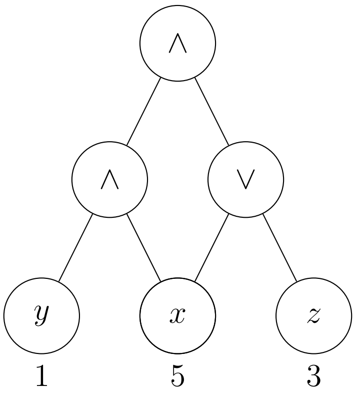

Figure 3.1: Attack tree with common variable

The Boolean function of the attack tree shown in a Figure 3.1 is $(yx)(x + z)$. The logical conjunction ($\wedge$) operator is marked as a multiplication and the disjunction ($\vee$) operator is marked as an addition. Therefore, $(yx)(x+z)$ means the same as $(y \wedge x) \wedge (x \vee z)$.

$$w(F) = \min\{w(yx), w(yxz)\} = \min\{6, 9\} = 6$$

An exact result for the function $F = (yx)(x+z)$ is 6.
Following the cost reduction method, the common variable $x$ is replaced by the two copies of it with reduced expenses $x'$, $x''$ so that:

$$w(x') + w(x'') + ... = w(x)$$

Figure 3.2 demonstrates that by dividing one common variable we get the exact result.

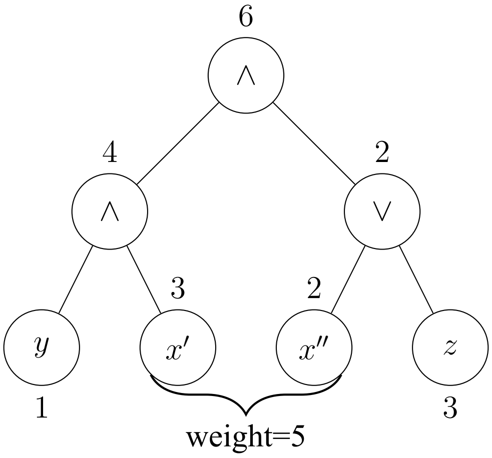

Figure 3.2: Exact result using cost reduction method ($x' = 3$ and $x'' = 2$)

Figure3.3 demonstrates another distribution of the common attack's cost $x$, where the result will not be exact when applying propagation method. In this case common variable $x$ is divided into variable $x'$ with weight 1 and $x''$ with weight 4.

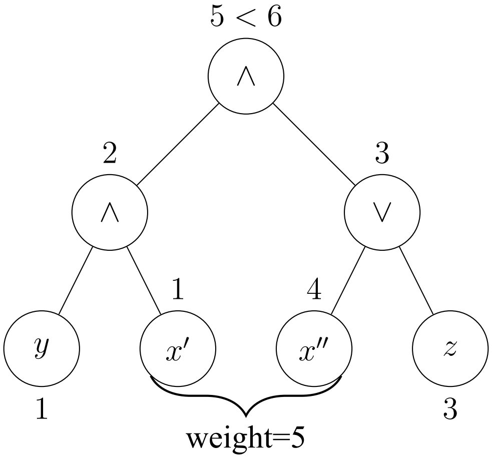

Figure 3.3: Non-exact result using cost reduction method ($x' = 1$ and $x'' = 4$)

The technique of artificial cost's reduction in its overall view was introduced in the articles of Buldas, A. and Lenin, A.[15, 2]. By using this technique, the result will not always be an exact value, as expenses taken for calculation can be less than they really are.

The articles [15, 2] also studied at the problem of protecting the system from rational attackers which got the name of Weighted Monotone Satisfiability (WMSAT). The principle includes the search for the answer about existence of solution $A$ for the monotone Boolean function $F(x_1, ..., x_n)$, whose weight would be less than the given limit ($w(A) < P$ ). In case such a solution does not exist, the system can be considered to be unattractive for attackers. Theorem 8 states that the Weighted Monotone Satisfiability problem is NP-complete. Typical NP-complete tasks are solved with the help of heuristic methods, which approximate the result. As it is known, heuristic methods can produce reasonably good solutions in most practical cases. Lenin A.[2] suggested to calculate the result using the method which approximates it ”from above” and with the help of propagation method the upper bound is calculated. The security analysis is based on analysing adversarial utility upper bounds. The lower bound is given on the basis of heuristics. In other words, the exact result will lie in an interval formed by the upper and lower bounds of adversarial utility.

As has been mentioned before, the propagation method gives the exact result only in the case when attack trees are independent.

# 4 Exact Cost Reduction

## 4.1 Current research

This study continues the research of Buldas A. and Lenin A., outlined in section 3.4 and focuses on more complex case, where there are 2 or more common attacks in the tree structure. The current research is focuses around the question of existence of symmetric or asymmetric division of 2 or more common attacks, where the cost reduction method will give an exact result. In other words, it is necessary to check hypothesis that *«for Boolean functions $F(x_1, ..., x_n, z_k, ..., z_l)$ and $G(y_1, ..., y_m, z_k, ..., z_l)$, where $z_k...z_l$ there are common variables of F and G and there is a suitable distribution of weights of common variables $\{\alpha, \beta, ...\}$ so that the propagation method will give the exact result».*

In order to prove that the hypothesis holds for every possible case, it is required to examine each possible case and it is almost impossible to do. Another way is to prove the correctness of the hypothesis with the help of the theorem. It is much easier to disprove the hypothesis. It is enough to find just one example in order to disprove the hypothesis. In the process of current research a counter-example has been found and it proves that the result may contain defect if we apply cost reduction to the function with 2 common variables. In Figure 4.1 there is an attack tree with 2 common attacks $u$ and $v$. Let the cost of each attack be equal to 1 €.

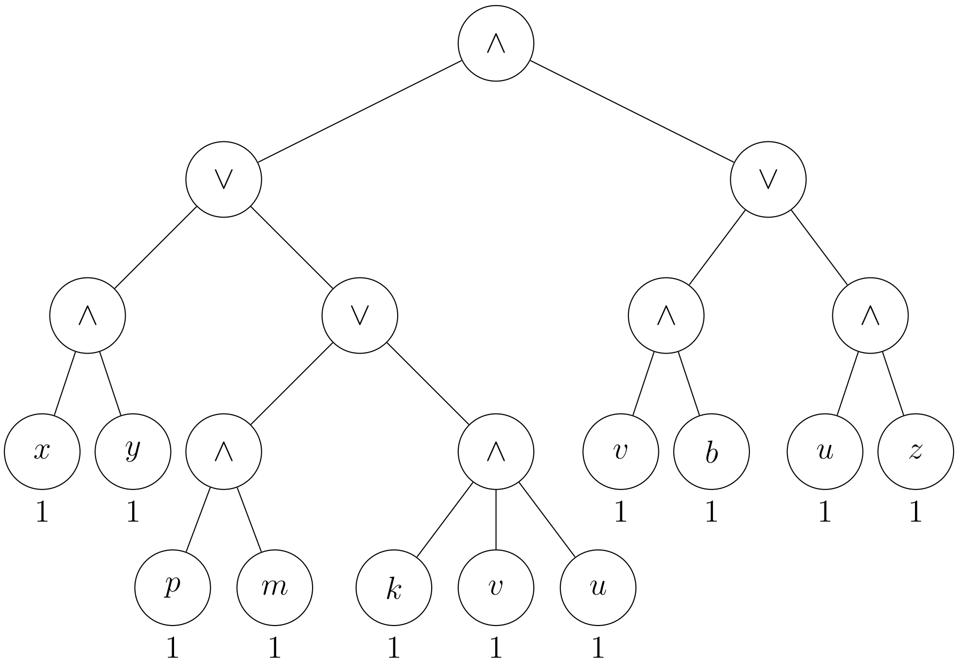

Figure 4.1: Attack tree with 2 common variables

Here and thereafter Boolean functions will be marked in the following way: logic operator conjunction ($\wedge$) is marked as multiplication ($\times$) and operator disjunction ($\vee$) is marked as addition (+). In order to find out the exact result of the function's solution it is necessary to choose the cheapest solution.

The Boolean function of the attack tree shown in Figure 4.1 is
$\phi = (xy + pm + vku)(vb + uz)$.
Therefore, $w(\phi) = \min\{w(xyvb), w(xyuz), w(pmvb), w(pmuz), w(vkub), w(vkuz)\} =$
$= \min\{4, 4, 4, 4, 4, 4\} = 4$ €
In the case of this example the value will be one of solutions which equals 4 €.
After implementing the cost reduction technique, the result will not be exact.

**Theorem 1:** There are functions $F$ and $G$ with 2 common variables $u$ and $v$ and an appropriate division of weight, that cost reduction method will always have a defect.
*Proof.* Let there be a Boolean function $\phi = FG$, where $F = (xy + pm + vku)$, and $G = (vb + uz)$. After dividing the variable $u$ into $u'$ and $u''$, $v$ into $v'$ and $v''$, the function $\phi$ takes the form:

$$\phi' = (xy + pm + v'ku')(v''b + u''z)$$

Therefore,
$w(\phi') = \min\{w(xyv''b), w(xyu''z), w(pmv''b), w(pmu''z), w(v'ku'v''b), w(v'ku'u''z)\} =$
$= \min\{w(xyv''b), w(xyu''z), w(pmv''b), w(pmu''z), w(vku'b), w(v'kuz)\} \leq$

$$\min\{w(xyv''b), w(v'kuz)\} = 3 + \min\{w(v'), w(v'')\} \leq 3.5 \quad \square$$

It is a complicated task to adapt and optimize the division of the corresponding weight of all common variables of the function. Depending on how we distribute weights, we will get different results. We focus on the best and worst weight distribution. If the difference is not big - such division of weight will be acceptable. If the interval between the results of the best and the worst distribution is big, it is necessary to optimize the division of weight in order to get more accurate approximation of the result.

The process of finding the result for NP-complete tasks is quite a hard process, as solving NP-complete tasks takes bigger amount of effort. Approximate values are sufficient for the purposes of the analysis. If we decide to solve an NP-complete task using cost reduction, we will not be able to get rid of the complexity of calculations, as components representing NP-complete task will appear in the calculations. It means that the calculation of the optimal division of the attack costs is a complicated task. If we apply heuristics (for instance symmetrical division), the result will be obtained in feasibly time, but it will be just an approximation of the precise result. The best and the worst distribution of weights common variables is shown in Figure 4.2.

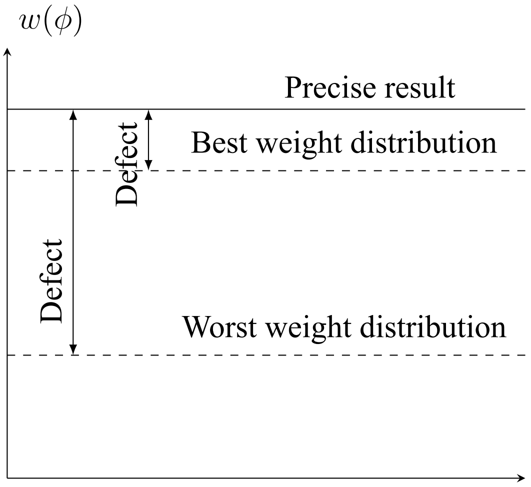

Figure 4.2: The best and the worst weight distribution

## 4.2 Notation

In the subsequent research we use the following setup. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a fixed set of independent variables. Let $\phi(x_1, x_2 \ldots, x_n)$ be a monotone Boolean function that depends on $x_1, x_2, \ldots, x_n$ but not on any other variables in $X$. Instead of the conjunction operator $\wedge$ we use an ordinary product notation, instead of the disjunction operator $\vee$ we use an ordinary sum notation. Thus, $x_1 \wedge x_2 \wedge x_3$ is written as $x_1x_2x_3$, similarly instead of $x_1 \vee x_2 \vee x_3$ is written as $x_1 + x_2 + x_3$. A *min-term* $\mu$ of $\phi$ is any conjunction $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ that implies truth of $\phi$. A weight function is any function $w : X \to \mathbb{R}^+$, which for any variable $x \in X$ returns a non-negative real number $w(x)$. If $\mu(x_1, x_2, \ldots, x_n) = x_1 x_2 \ldots x_n$, then $w(\mu) = w(x_1) + w(x_2) + \ldots + w(x_n)$. If $\phi$ is any Boolean function, then $w(\phi) = \min\{w(\mu) : \mu \text{ is a min-term of } \phi\}$.

Let $F(x_1, \ldots, x_n, z)$ and $G(y_1, \ldots, y_m, z)$ be monotone Boolean functions. The variable $z$ is the only common variable in $F$ and $G$. Let $F'(x_1, \ldots, x_n, z')$ and $G'(y_1, \ldots, y_m, z'')$ be the modified Boolean functions where $z$ has been replaced with its independent copies $z'$ and $z''$ respectively. $F'$ and $G'$ have no common variables, and $w(z') + w(z'') = w(z)$.

At this stage of research we focus on the case where conjunction which contains common attacks is a binary operator.

## 4.3 Optimal cost reduction

This chapter focuses on the reduction defect that may be present in the case of optimal distribution of weights of common variables.
Let there be a Boolean function $\phi = FG$ with common variable $z$.
The function $F$ can be represented as:

$$F = F_0 + z \cdot \partial F,$$

where $F_0, \partial F$ - parts of $F$ which do not depend on $z$, and $\partial F$ is so-called Boolean derivative ($\partial F = \frac{\partial}{\partial z} F$).
The function $G$ can be represented as:

$$G = G_0 + z \cdot \partial G,$$

where $G_0, \partial G$ - parts of $G$ that do not depend on $z$, and $\partial G$ is so-called Boolean derivative ($\partial G = \frac{\partial}{\partial z} G$).

**Theorem 2:** The reduction defect is non-zero ($\triangle > 0$) if the following inequalities hold

$$|w(G_0 \partial F) - w(F_0 \partial G)| < w(z) < 2m_0 - w(F_0 \partial G) - w(G_0 \partial F, ) \tag{4.1}$$

and the reduction defect has the following upper bound:

$$\triangle \le \frac{w(z)}{2} - \frac{1}{2}|w(G_0 \partial F) - w(F_0 \partial G)| \le \frac{w(z)}{2} \tag{4.2}$$

*Proof.* Let $\phi = FG$ be a Boolean function and $z$ be the common variable of $F$ and $G$. We can represent the function $\phi$ in the following way:

$$FG = (F_0 + z\partial F)(G_0 + z\partial G) = F_0 G_0 + G_0 z \partial F + F_0 z \partial G + z \partial G \partial F$$

Having divided variable $z$ to independent variables $z'$ and $z''$, we get a new function:

$$F'G' = (F_0 + z'\partial F)(G_0 + z''\partial G) = F_0G_0 + G_0z'\partial F + F_0z''\partial G + \underbrace{z'z''\partial G\partial F}_{z\partial G\partial F}$$

It is known that $w(z) = w(z') + w(z'')$, therefore:

$$w(z'z''\partial F\partial G) = w(z\partial F\partial G)$$

Reduction defect $\triangle$ is the difference between $w(FG)$ and $w(F'G')$.
$\triangle = w(FG) - \max\limits_{\alpha} w(F'G')$, where $\triangle$ is not a negative value as according to Theorem. 4.5.5 [2] $w(FG) \geq w(F'G')$.
The value of primary and transformed function is calculated by formulas introduced further:

$$w(FG) = \min\{\underbrace{w(F_0G_0), w(z\partial F\partial G)}_{m_0}, w(zG_0\partial F), w(zF_0\partial G)\},$$

$$w(F'G') = \max_{\alpha}\min\{\underbrace{w(F_0G_0), w(z'z''\partial F\partial G)}_{m_0}, w(z'G_0\partial F), w(z''F_0\partial G)\},$$

$$m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\},$$

where $w(z) = w(z') + w(z'')$
The defect is calculated using the formula:
$\triangle = w(FG) - \max\limits_{\alpha} w(F'G') =$
$= \overbrace{\min\{w(m_0), w(zG_0\partial F), w(zF_0\partial G)\}}^{w(FG)} - \max\limits_{\alpha} \overbrace{\min\{w(m_0), w(z'G_0\partial F), w(z''F_0\partial G)\}}^{w(F'G')}.$
Defect will be positive ($\triangle \geq 0$), if

$$\min\{w(m_0), w(zG_0\partial F), w(zF_0\partial G)\} > \max_{\alpha}\min\{w(m_0), w(z'G_0\partial F), w(z''F_0\partial G)\} \tag{4.3}$$

2 conditions can be formulated:

$$\min\{zG_0\partial F, zF_0\partial G\} > \max_{\alpha}\min\{z'G_0\partial F, z''F_0\partial G\} \tag{4.4}$$

$$m_0 > \max_{\alpha}\min\{z'G_0\partial F, z''F_0\partial G\} \tag{4.5}$$

2 cases should be considered here:

A) $w(G_0\partial F) \geq w(F_0\partial G)$
B) $w(G_0\partial F) \leq w(F_0\partial G)$

Case A assumes, that $w(zF_0\partial G)$ will be less or equal to $w(zG_0\partial F)$. Therefore, $w(zF_0\partial G)$ is the minimal value in $\min\{w(zG_0\partial F), w(zF_0\partial G)\}$.
Inequalities (4.4) and (4.5) can be re-written as follows:

$$\max_{\alpha}\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} < w(zF_0\partial G) \tag{4.6}$$

$$\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} < m_0 \tag{4.7}$$

It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\}$ will reach its maximum when $w(z'G_0\partial F) = w(z''F_0\partial G)$, as shown in the Figure 4.3, where $w(z') = \alpha \cdot w(z)$ and $w(z'') = (1-\alpha) \cdot w(z)$.

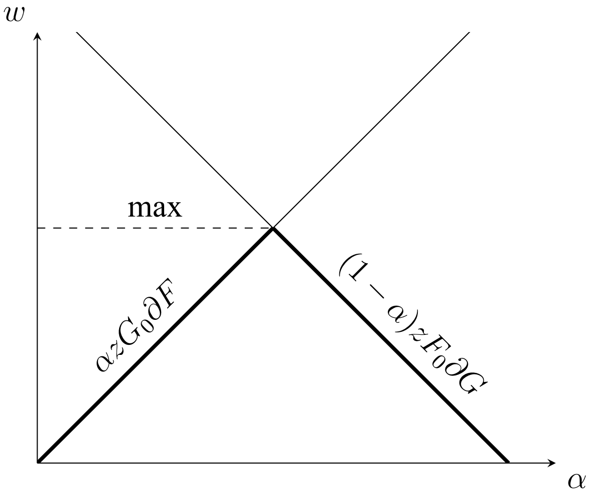

Figure 4.3: Maximums of function $\max_{\alpha} \min\{w(\alpha z G_0\partial F), w((1-\alpha)zF_0\partial G)\}$

$w(z'G_0\partial F) = w(z''F_0\partial G)$
$w(z') + w(G_0\partial F) = w(z) - w(z') + w(F_0\partial G)$
$w(z') + w(z') = w(z) + w(F_0\partial G) - w(G_0\partial F)$
$2w(z') = w(z) + w(F_0\partial G) - w(G_0\partial F)$

$$w(z') = \frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) - w(G_0\partial F)\right] \tag{4.8}$$

Considering case A, equation (4.8) is transformed into:
$w(z') = \frac{w(z)}{2} - \frac{1}{2}\left[w(G_0\partial F) - w(F_0\partial G)\right] = \frac{w(z)}{2} - \frac{1}{2}|w(G_0\partial F) - w(F_0\partial G)|$
Therefore we have obtained the formula for the optimal value of $w(z')$.

According to (4.3):
$\min\{w(m_0), w(zG_0\partial F), w(zF_0\partial G)\} > \max_{\alpha}\{w(m_0), w(z'G_0\partial F), w(z''F_0\partial G)\}$.
It is also known that the maximal value is achieved in the case when $w(z'G_0\partial F) = w(z''F_0\partial G)$. Therefore it means that any of the 2 values can be chosen for calculation. So $\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = w(z'G_0\partial F)$.
The value $w(z'G_0\partial F)$ can be re-written as the sum of $w(z')$ and $w(G_0\partial F)$. Instead of $w(z')$ we can use the value achieved from previous calculations. Transformations look as follows:
$\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = w(z'G_0\partial F) = w(z') + w(G_0\partial F) =$
$= \frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) - w(G_0\partial F)\right] + w(G_0\partial F) =$

$= \frac{w(z)}{2} + \frac{1}{2}w(F_0\partial G) - \frac{1}{2}w(G_0\partial F) + w(G_0\partial F) =$
$= \frac{w(z)}{2} + \frac{1}{2}w(F_0\partial G) + \frac{1}{2}w(G_0\partial F) = \frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right]$

The left-hand side of inequality (4.6) can be replaced by $\frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right]$.
$\frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right] < w(zF_0\partial G)$
$\frac{w(z)}{2} + \frac{1}{2}w(F_0\partial G) + \frac{1}{2}w(G_0\partial F) < w(z) + w(F_0\partial G)$.
We simplify this inequality by getting rid of the fractional parts and then multiply both parts of the condition by 2.
$w(z) + w(F_0\partial G) + w(G_0\partial F) < 2w(z) + 2w(F_0\partial G)$
$w(G_0\partial F) < 2w(z) + 2w(F_0\partial G) - w(z) - w(F_0\partial G)$
$w(G_0\partial F) < w(z) + w(F_0\partial G)$
$w(G_0\partial F) - w(F_0\partial G) < w(z)$
Therefore, $w(z) > w(G_0\partial F) - w(F_0\partial G)$, which corresponds to

$$w(z) > |w(G_0\partial F) - w(F_0\partial G)|, \tag{4.9}$$

given case A holds.

The left-hand side of inequality (4.7) can be replaced by $\frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right]$.
$\frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right] < m_0$
We multiply both parts of inequality by 2.
$w(z) + w(F_0\partial G) + w(G_0\partial F) < 2m_0$
$w(z) < 2m_0 - w(F_0\partial G) - w(G_0\partial F)$
Therefore,

$$w(z) < 2m_0 - w(F_0\partial G) - w(G_0\partial F) \tag{4.10}$$

Inequalities (4.9) and (4.10) can be joined together into
$|w(G_0\partial F) - w(F_0\partial G)| < w(z) < 2m_0 - w(F_0\partial G) - w(G_0\partial F)$, which corresponds to (4.1).

$$\triangle = \min\{m_0, zG_0\partial F, zF_0\partial G\} - \max_{\alpha} \min\{m_0, z'G_0\partial F, z''F_0\partial G\} \leq$$
$$\leq \min\{zG_0\partial F, zF_0\partial G\} - \max_{\alpha} \min\{z'G_0\partial F, z''F_0\partial G\}$$

As for considered case where $w(zF_0\partial G) < w(zG_0\partial F)$, the minimal value is $w(zF_0\partial G)$, let us substitute the appropriate value in the formula.

$$\triangle = w(zF_0\partial G) - \left(\frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right]\right)$$

$\triangle = w(zF_0\partial G) - \left(\frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) + w(G_0\partial F)\right]\right) =$
$= w(zF_0\partial G) - \frac{w(z)}{2} - \frac{1}{2}w(F_0\partial G) - \frac{1}{2}w(G_0\partial F) =$

$= w(z) + w(F_0\partial G) - \frac{w(z)}{2} - \frac{1}{2}w(F_0\partial G) - \frac{1}{2}w(G_0\partial F) =$
$= \frac{w(z)}{2} + \frac{w(F_0\partial G)}{2} - \frac{w(G_0\partial F)}{2} = \frac{w(z)}{2} + \frac{1}{2}\left[w(F_0\partial G) - w(G_0\partial F)\right] =$
$= \frac{w(z)}{2} - \frac{1}{2}\left[w(G_0\partial F) - w(F_0\partial G)\right] \leq \frac{w(z)}{2}$, that is identical to
$\frac{w(z)}{2} - \frac{1}{2}|w(G_0\partial F) - w(F_0\partial G)| \leq \frac{w(z)}{2}$ given case A holds.
Therefore the defect is calculated using the following formula:
$\triangle = \frac{w(z)}{2} - \frac{1}{2}|w(G_0\partial F) - w(F_0\partial G)| \leq \frac{w(z)}{2}$, which corresponds to (4.2).

Case B assumes, that $w(zG_0\partial F)$ will be less or equal to $w(zF_0\partial G)$. Therefore, $w(zG_0\partial F)$ is the minimal value in $\min\{w(zG_0\partial F), w(zF_0\partial G)\}$.
Inequalities (4.4) and (4.5) can be re-written as follows:

$$\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} < w(zG_0\partial F) \tag{4.11}$$

$$\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} < m_0 \tag{4.12}$$

It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\}$ will reach its maximum when $w(z'G_0\partial F) = w(z''F_0\partial G)$, as shown in the Figure 4.3.

$w(z'G_0\partial F) = w(z''F_0\partial G)$
$w(z) - w(z'') + w(G_0\partial F) = w(z'') + w(F_0\partial G)$
$w(z) + w(G_0\partial F) - w(F_0\partial G) = w(z'') + w(z'')$
$2w(z'') = w(z) + w(G_0\partial F) - w(F_0\partial G)$

$$w(z'') = \frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) - w(F_0\partial G)\right] \tag{4.13}$$

Following case B, the equation (4.13) is transformed into:
$w(z'') = \frac{w(z)}{2} - \frac{1}{2}\left[w(F_0\partial G) - w(G_0\partial F)\right] = \frac{w(z)}{2} - \frac{1}{2}|w(F_0\partial G) - w(G_0\partial F)|$
Therefore we have obtained the formula for the optimal value of $w(z'')$.
According to (4.3):
$\min\{w(m_0), w(zG_0\partial F), w(zF_0\partial G)\} > \max_{\alpha}\{w(m_0), w(z'G_0\partial F), w(z''F_0\partial G)\}$.
It is also known that the maximal value is achieved in the case when $w(z'G_0\partial F) = w(z''F_0\partial G)$. Therefore it means that any of the 2 values can be chosen for calculation. So $\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = w(z''F_0\partial G)$.
The value $w(z''F_0\partial G)$ can be re-written as the sum of $w(z'')$ and $w(F_0\partial G)$. Instead of $w(z'')$ we can use the value achieved in previous calculations. Transformations look this way:

$\max_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = w(z''F_0\partial G) = w(z'') + w(F_0\partial G) =$
$= \frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) - w(F_0\partial G)\right] + w(F_0\partial G) =$

$= \frac{w(z)}{2} + \frac{1}{2}w(G_0\partial F) - \frac{1}{2}w(F_0\partial G) + w(F_0\partial G) =$
$= \frac{w(z)}{2} + \frac{1}{2}w(G_0\partial F) + \frac{1}{2}w(F_0\partial G) = \frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) + w(F_0\partial G)\right]$

The left-hand side of inequality (4.11) can be replaced by $\frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) + w(F_0\partial G)\right]$.
$\frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) + w(F_0\partial G)\right] < w(zG_0\partial F)$
$\frac{w(z)}{2} + \frac{1}{2}w(G_0\partial F) + \frac{1}{2}w(F_0\partial G) < w(z) + w(G_0\partial F)$.
We simplify this inequality by getting rid of the fractional parts and then multiply both parts of the condition by 2.
$w(z) + w(G_0\partial F) + w(F_0\partial G) < 2w(z) + 2w(G_0\partial F)$
$w(F_0\partial G) < 2w(z) + 2w(G_0\partial G) - w(z) - w(G_0\partial F)$
$w(F_0\partial G) < w(z) + w(G_0\partial F)$
$w(F_0\partial G) - w(G_0\partial F) < w(z)$
Therefore, $w(z) > w(F_0\partial G) - w(G_0\partial F)$, that is identical to

$$w(z) > |w(F_0\partial G) - w(G_0\partial F)|, \tag{4.14}$$

given case B holds.

The left-hand side of inequality (4.12) can be replaced by $\frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial G) + w(F_0\partial F)\right]$.
$\frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) + w(F_0\partial G)\right] < m_0$
We multiply both parts of inequality by 2.
$w(z) + w(G_0\partial F) + w(F_0\partial G) < 2m_0$
$w(z) < 2m_0 - w(G_0\partial F) - w(F_0\partial G)$
Therefore,

$$w(z) < 2m_0 - w(G_0\partial F) - w(F_0\partial G) \tag{4.15}$$

Inequalities (4.14) and (4.15) can be joined together into
$|w(F_0\partial G) - w(G_0\partial F)| < w(z) < 2m_0 - w(G_0\partial F) - w(F_0\partial G)$, which corresponds to (4.1).

$$\triangle = \min\{m_0, zG_0\partial F, zF_0\partial G\} - \max_{\alpha} \min\{m_0, z'G_0\partial F, z''F_0\partial G\} \leq$$
$$\leq \min\{zG_0\partial F, zF_0\partial G\} - \max_{\alpha} \min\{z'G_0\partial F, z''F_0\partial G\}$$

As for considered case where $w(zG_0\partial F) < w(zF_0\partial G)$, the minimum value is $w(zG_0\partial F)$, let us substitute the appropriate value to the formula.

$$\triangle = w(zG_0\partial F) - \left(\frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) + w(F_0\partial G)\right]\right)$$

$\triangle = w(zG_0\partial F) - \left(\frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) + w(F_0\partial G)\right]\right) =$
$= w(zG_0\partial F) - \frac{w(z)}{2} - \frac{1}{2}w(G_0\partial F) - \frac{1}{2}w(F_0\partial G) =$

$= w(z) + w(G_0\partial F) - \frac{w(z)}{2} - \frac{1}{2}w(G_0\partial F) - \frac{1}{2}w(F_0\partial G) =$
$= \frac{w(z)}{2} + \frac{w(G_0\partial F)}{2} - \frac{w(F_0\partial G)}{2} = \frac{w(z)}{2} + \frac{1}{2}\left[w(G_0\partial F) - w(F_0\partial G)\right] =$
$= \frac{w(z)}{2} - \frac{1}{2}\left[w(F_0\partial G) - w(G_0\partial F)\right] \leq \frac{w(z)}{2}$, that is identical to
$\frac{w(z)}{2} - \frac{1}{2}|w(F_0\partial G) - w(G_0\partial F)| \leq \frac{w(z)}{2}$ given case B holds.
Therefore the defect is calculated according to the formula:
$\triangle = \frac{w(z)}{2} - \frac{1}{2}|w(F_0\partial G) - w(G_0\partial F)| \leq \frac{w(z)}{2}$, that corresponds to (4.2). □

## 4.4 Worst cost reduction

The current chapter will have an assessment of the defect which can be received in the process of division of 2 common variables with the worst weight's division of common variables.
Let there be a Boolean function $\phi = FG$ with common variable $z$. Functions $F$ and $G$ can be represented as:

$$F = F_0 + z \cdot \partial F,$$

$$G = G_0 + z \cdot \partial G,$$

where $F_0, \partial F, G_0, \partial G$ are the parts of the functions which do not depend on $z$. The function $\partial F$ is the so-called Boolean derivative $\partial F = \frac{\partial}{\partial z}F$ and the function $\partial G$ is the so-called Boolean derivative $\partial G = \frac{\partial}{\partial z}G$.

**Theorem 3:** The reduction defect is non-zero ($\triangle > 0$) if either of the following inequalities hold:

$$\begin{cases} w(G_0\partial F) < w(z) + w(F_0\partial G), \\ w(G_0\partial F) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\} \end{cases} \tag{4.16}$$

or

$$\begin{cases} w(F_0\partial G) < w(z) + w(G_0\partial F), \\ w(F_0\partial G) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\} \end{cases} \tag{4.17}$$

and has the following upper bound:

$$\triangle \leq w(z) - |w(F_0\partial G) - w(G_0\partial F)| \tag{4.18}$$

*Proof.* Let $\phi = FG$ be a Boolean function and $z$ be the common variable of $F$ and $G$.
We can represent the function $\phi$ in the following way:

$$FG = (F_0 + z\partial F)(G_0 + z\partial G) = F_0G_0 + G_0z\partial F + F_0z\partial G + z\partial G\partial F$$

Having divided variable $z$ into independent variables $z'$ and $z''$, functions $F'$ and $G'$ can be shown as:

$$F' = F_0 + z' \cdot \partial F$$

$$G' = G_0 + z'' \cdot \partial G$$

The modified Boolean function $\phi$ can be re-written as follows:

$$F'G' = (F_0 + z'\partial F)(G_0 + z''\partial G) = F_0G_0 + G_0z'\partial F + F_0z''\partial G + \underbrace{z'z''\partial G\partial F}_{z\partial G\partial F}$$

Reduction defect $\triangle$ is the difference between $w(FG)$ and $w(F'G')$.

The values of the initial and transformed functions can be calculated using formulas introduced further:

$$w(FG) = \min\{\underbrace{w(F_0G_0), w(z\partial F\partial G)}_{m_0}, w(zG_0\partial F), w(zF_0\partial G)\}$$

$$w(F'G') = \max_{\alpha} \min\{\underbrace{w(F_0G_0), w(z'z''\partial F\partial G)}_{m_0}, w(z'G_0\partial F), w(z''F_0\partial G)\},$$

$$m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\},$$

where $w(z) = w(z') + w(z'')$

Defect is positive ($\triangle \geq 0$), if:

$$\min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} < \min\{w(zG_0\partial F), w(zF_0\partial G)\} \tag{4.19}$$

$$\min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\} \tag{4.20}$$

$\min\{w(z'G_0\partial F), w(z''F_0\partial G)\}$ will reach its minimum when $w(z') = 0$ or $w(z') = w(z)$, as shown in the Figure 4.4.

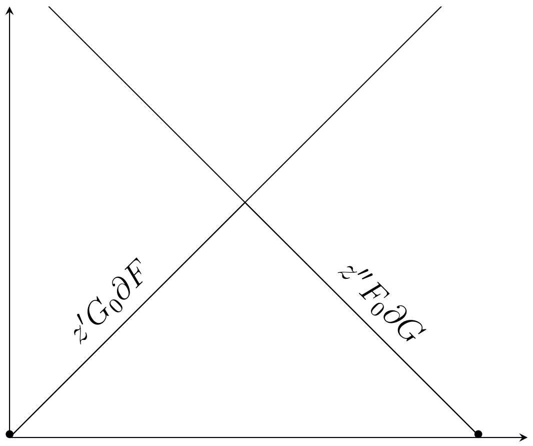

Figure 4.4: Minimums of function $\min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\}$

It is known that

$$w(z) = w(z') + w(z'') \tag{4.21}$$

If $w(z') = 0$, then according to (4.21) $w(z'') = w(z) - w(z') = w(z) - 0 = w(z)$. Considering, that $w(z') = 0$ and $w(z'') = w(z)$, the left-hand side of inequality (4.19) can be replaced as follows:
$\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(G_0\partial F), w(zF_0\partial G)\} =$
$= \min\{w(G_0\partial F), w(z) + w(F_0\partial G)\}$.

$$\min_{\alpha} \min\{w(G_0\partial F), w(z) + w(F_0\partial G)\} < min\{w(zG_0\partial F), w(zF_0\partial G)\}$$

3 cases should be considered here:

$$\text{A) } w(G_0\partial F) > w(z) + w(F_0\partial G) > w(F_0\partial G)$$
$$\text{B) } w(F_0\partial G) < w(G_0\partial F) < w(z) + w(F_0\partial G)$$
$$\text{C) } w(F_0\partial G) > w(G_0\partial F)$$

Case A assumes, that $w(G_0\partial F) > w(z) + w(F_0\partial G) > w(F_0\partial G)$.
It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(G_0\partial F), w(z) + w(F_0\partial G)\}$.
Then, $\min\{w(G_0\partial F), w(z) + w(F_0\partial G)\} = w(z) + w(F_0\partial G)$ given case A holds.

$\min\{w(zG_0\partial F), w(zF_0\partial G)\} = w(zF_0\partial G) = w(z) + w(F_0\partial G)$

Inequalities (4.19) and (4.20) can be re-written as follows:

$$w(z) + w(F_0\partial G) < w(z) + w(F_0\partial G)$$

$$w(z) + w(F_0\partial G) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\}$$

It can be seen that the first inequality never holds.

Case B assumes, that $w(F_0\partial G) < w(G_0\partial F) < w(z) + w(F_0\partial G)$. It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(G_0\partial F), w(z) + w(F_0\partial G)\}$.
Then, $\min\{w(G_0\partial F), w(z) + w(F_0\partial G)\} = w(G_0\partial F)$ given case B holds.

$\min\{w(zG_0\partial F), w(zF_0\partial G)\} = w(zF_0\partial G) = w(z) + w(F_0\partial G)$

Inequalities (4.19) and (4.20) can be re-written as follows:

$$w(G_0\partial F) < w(z) + w(F_0\partial G)$$

$$w(G_0\partial F) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\}$$

The defect is calculated by the following formula:
$\triangle = w(FG) - \min_{\alpha} w(F'G') =$
$= \min\{m_0, w(zG_0\partial F), w(zF_0\partial G)\} - \min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} \leq$
$\leq \min\{w(zG_0\partial F), w(zF_0\partial G)\} - \min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} =$
$= w(z)+w(F_0\partial G)-w(G_0\partial F) = w(z)-[w(G_0\partial F) - w(F_0\partial G)] = w(z)-|w(G_0\partial F)- w(F_0\partial G)|$, that corresponds to (4.18), according to B.

Case C assumes, that $w(F_0\partial G) > w(G_0\partial F)$.
It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(G_0\partial F), w(z) + w(F_0\partial G)\}$.
Then, $\min\{w(G_0\partial F), w(z)+w(F_0\partial G)\} = w(G_0\partial F)$, as $w(z)+w(F_0\partial G) > w(G_0\partial F)$, given case C holds.

$\min\{w(zG_0\partial F), w(zF_0\partial G)\} = w(zG_0\partial F) = w(z) + w(G_0\partial F)$

Inequalities (4.19) and (4.20) can be re-written as follows:

$$w(G_0\partial F) < w(z) + w(G_0\partial F)$$

$$w(G_0\partial F) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\}$$

The defect is calculated by the this formula:
$\triangle = w(FG) - \min_{\alpha} w(F'G') =$
$= \min\{m_0, w(zG_0\partial F), w(zF_0\partial G)\} - \min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} \leq$
$\leq \min\{w(zG_0\partial F), w(zF_0\partial G)\} - \min_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} =$
$= w(z) + w(G_0\partial F) - w(G_0\partial F) = w(z)$, that corresponds to (4.18), according to C.

If $w(z') = w(z)$, then according to (4.21) $w(z'') = w(z) - w(z') = w(z) - w(z) = 0$.
Considering that $w(z') = w(z)$ and $w(z'') = 0$, the left-hand side of inequality (4.19) can be replaced as follows:
$\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(z) + w(G_0\partial F), w(F_0\partial G)\}$.

3 cases should be considered here:

D) $w(F_0\partial F) > w(z) + w(G_0\partial F) > w(G_0\partial F)$
E) $w(G_0\partial F) < w(F_0\partial G) < w(z) + w(G_0\partial F)$
F) $w(G_0\partial F) > w(F_0\partial G)$

Case D assumes, that $w(F_0\partial F) > w(z) + w(G_0\partial F) > w(G_0\partial F)$.
It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(z) + w(G_0\partial F), w(F_0\partial G)\}$.

Then, $\min\{w(z)+w(G_0\partial F), w(F_0\partial G)\} = w(z)+w(G_0\partial F)$, given case D holds.

$\min\{w(zG_0\partial F), w(zF_0\partial G)\} = w(zG_0\partial F) = w(z)+w(G_0\partial F)$

Inequalities (4.19) and (4.20) can be re-written as follows:

$$w(z)+w(G_0\partial F) < w(z)+w(G_0\partial F)$$

$$w(z)+w(G_0\partial F) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\}$$

It can be seen that the first inequality never holds.

Case E assumes, that $w(G_0\partial F) < w(F_0\partial G) < w(z)+w(G_0\partial F)$. It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(z)+w(G_0\partial F), w(F_0\partial G)\}$.
Then, $\min\{w(z)+w(G_0\partial F), w(F_0\partial G)\} = w(F_0\partial G)$ given case E holds.

$\min\{w(zG_0\partial F), w(zF_0\partial G)\} = w(zG_0\partial F) = w(z)+w(G_0\partial F)$

Inequalities (4.19) and (4.20) can be re-written as follows:

$$w(F_0\partial G) < w(z)+w(G_0\partial F)$$

$$w(F_0\partial G) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\}$$

The defect is calculated by the following formula:
$\triangle = w(FG) - \min\limits_{\alpha} w(F'G') =$
$= \min\{m_0, w(zG_0\partial F), w(zF_0\partial G)\} - \min\limits_{\alpha}\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} \leq$
$\leq \min\{w(zG_0\partial F), w(zF_0\partial G)\} - \min\limits_{\alpha}\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} =$
$= w(z)+w(G_0\partial F)-w(F_0\partial G) = w(z)-[w(F_0\partial G)-w(G_0\partial F)] = w(z)-|w(F_0\partial G)-w(G_0\partial F)|$, that corresponds to (4.18), according to E.

Case F assumes, that $w(G_0\partial F) > w(F_0\partial G)$.
It is known that $\min\{w(z'G_0\partial F), w(z''F_0\partial G)\} = \min\{w(z)+w(G_0\partial F), w(F_0\partial G)\}$.
Then, $\min\{w(z)+w(G_0\partial F), w(F_0\partial G)\} = w(F_0\partial G)$ given case F holds.

$\min\{w(zG_0\partial F), w(zF_0\partial G)\} = w(zF_0\partial G) = w(z)+w(F_0\partial G)$

Inequalities (4.19) and (4.20) can be re-written as follows:

$$w(F_0\partial G) < w(z)+w(F_0\partial G)$$

$$w(F_0\partial G) < m_0 = \min\{w(F_0G_0), w(z\partial F\partial G)\}$$

The defect is calculated by the this formula:
$\triangle = w(FG) - \min\limits_{\alpha} w(F'G') =$
$= \min\{m_0, w(zG_0\partial F), w(zF_0\partial G)\} - \min\limits_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} \leq$
$\leq \min\{w(zG_0\partial F), w(zF_0\partial G)\} - \min\limits_{\alpha} \min\{w(z'G_0\partial F), w(z''F_0\partial G)\} =$
$= w(z) + w(F_0\partial G) - w(F_0\partial G) = w(z)$, that corresponds to (4.18), according to F. □

By dividing the weight of a single common variable $z$ in the best possible way, the defect can reach the size which equals $\frac{w(z)}{2}$. In the worst case where division of each variable will provide the defect, the common defect will equal the half of the sum of all common variables' weights. In the case of most inefficient reduction of common variable $z$ the defect will reach the size which equals $w(z)$. In the worst case where the division of each variable will give defect, the common defect will equal the weight's sum of all variables.

The difference between the best and the worst result of optimization equals $\triangle = \mid \triangle_b - \triangle_w \mid = |\frac{w(z_1)+\ldots+w(z_n)}{2} - w(z_1) - \ldots - w(z_n)| = \frac{w(z_1)+\ldots+w(z_n)}{2}$, where $\triangle_b$ is the defect in the best weight's division, $\triangle_w$ is the defect with the worst weight's division and $x_1, x_2, \ldots x_n$ are the common variables. This is illustrated in Figure 4.5.

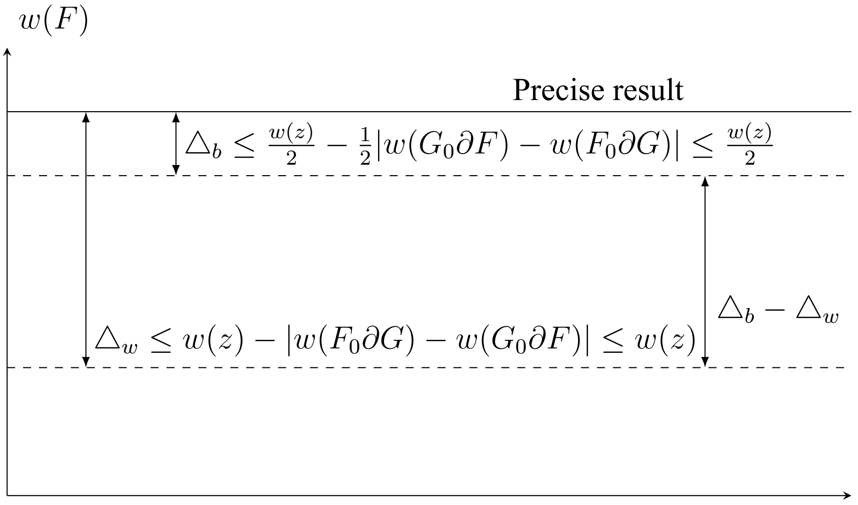
$w(F)$
Precise result
$\triangle_b \leq \frac{w(z)}{2} - \frac{1}{2}|w(G_0\partial F) - w(F_0\partial G)| \leq \frac{w(z)}{2}$
$\triangle_b - \triangle_w$
$\triangle_w \leq w(z) - |w(F_0\partial G) - w(G_0\partial F)| \leq w(z)$

Figure 4.5: Defect for different weight distribution

On the basis of the obtained results it is possible to make the following conclusions. If the number of common attacks is not high and their cost is not expensive, it is possible to sacrifice the accuracy of the result and to take any values in division for the weights of common variables. For instance, symmetrical division of common variables can be used. Less time is needed with such approach rather than with optimization of attacks' cost division. The reduction defect will not be big in this case. If the number of common attacks is high or the costs are expensive, the size of the reduction defect, reducing costs in an occasional way, may be high, which is not acceptable. In this case it makes sense to invest a bit more resources and optimize the attack cost division in a way when the values of divided variables become the best one. The defect will be minimal in this case.

## 4.5 Proof of concept

In order to illustrate previously shown theorems we can give an example. Formulas from Theorem 2 and Theorem 3 will be used.

It is known that in division of the first variable $v$:

$$F = (xy + pm + vku) = \underbrace{(xy + pm)}_{F_0} + v \cdot \underbrace{(ku)}_{\partial F},$$

$$G = (vb + uz) = \underbrace{(uz)}_{G_0} + v \cdot \underbrace{b}_{\partial G}$$

In division of the 2-nd common variable $u$:

$$F = (xy + pm + v'ku) = \underbrace{(xy + pm)}_{F_0} + u \cdot \underbrace{(kv')}_{\partial F},$$

$$G = (v''b + uz) = \underbrace{(v''b)}_{G_0} + u \cdot \underbrace{z}_{\partial G}$$

Let us divide the weight of the first common variable $v$.
$w(F_0) = w(xy) + w(pm) = 2$
$w(\partial F) = w(ku) = 2$
$w(G_0) = w(uz) = 2$
$w(\partial G) = w(b) = 1$
$w(G_0\partial F) = w(uzku) = uzk = 3$
$w(F_0\partial G) = w((xy + pm)b) = 2 + 1 = 3$
$w(F_0G_0) = w((xy + pm)uz) = 2 + 2 = 4$
$w(v\partial F\partial G) = w(vkub) = 4$
$w(m_0) = \min\{w(F_0G_0), w(v\partial F\partial G)\} = \min\{4, 4\} = 4$
$2w(m_0) = 2 \cdot 4 = 8$
By using the best cost reduction we check if the condition is completed:

$$|w(G_0\partial F) - w(F_0\partial G)| < w(v) < 2m_0 - w(F_0\partial G) - w(G_0\partial F)$$

$$|3 - 3| < 1 < 8 - 3 - 3$$

$$0 < 1 < 2$$

According to the results of calculations the condition is evaluated to true, therefore in division of the common variable's weight $v$ in the best way, the defect $\triangle$ will exist. Let us calculate the defect with the best weight division.

$w(v) = 1$

$w(G_0 \partial F) = w(uzku) = uzk = 3$

$w(F_0 \partial G) = w((xy + pm)b) = 2 + 1 = 3$

$$\triangle = \frac{w(v)}{2} - \frac{1}{2} | \underbrace{w(G_0 \partial F) - w(F_0 \partial G)}_{0} | \leq \frac{w(v)}{2}$$

$$\triangle = \frac{1}{2} - \frac{1}{2} |3 - 3| \leq \frac{1}{2}$$

$$\triangle = \frac{1}{2} - \frac{1}{2} \cdot 0 \leq \frac{1}{2}$$

$$\triangle = \frac{1}{2} \leq \frac{1}{2}$$

By using the worst cost reduction we check if the conditions are completed:

$$\begin{cases} w(G_0 \partial F) < w(v) + w(F_0 \partial G), \\ w(G_0 \partial F) < m_0 = \text{min}\{w(F_0 G_0), w(v \partial F \partial G)\} \end{cases}$$

or

$$\begin{cases} w(F_0 \partial G) < w(v) + w(G_0 \partial F), \\ w(F_0 \partial G) < m_0 = \text{min}\{w(F_0 G_0), w(v \partial F \partial G)\} \end{cases}$$

$$\begin{cases} 3 < 1 + 3, \\ 3 < 4 \end{cases} \quad \begin{cases} 3 < 4, \\ 3 < 4 \end{cases}$$

or

$$\begin{cases} 3 < 1 + 3, \\ 3 < 4 \end{cases} \quad \begin{cases} 3 < 4, \\ 3 < 4 \end{cases}$$

According to the results, the condition is evaluated to true, therefore in division of the common variable's weight $v$ in the worst way, the defect will exist. Let us calculate the defect $\triangle$ with the worst weight division:

$w(v) = 1$

$w(G_0 \partial F) = w(uzku) = uzk = 3$

$w(F_0 \partial G) = w((xy + pm)b) = 2 + 1 = 3$

$$\triangle \leq w(v) - |w(F_0 \partial G) - w(G_0 \partial F)|$$

$$\triangle \leq 1 - |3 - 3|$$

$$\triangle \leq 1 - 0$$

$$\triangle \leq 1$$

Let us divide the cost of the second variable $u$.

In order to apply cost reduction method to the second common variable, it is necessary to calculate the optimal values of the variables $v'$ and $v''$. In the best optimization the optimal values $v'$ and $v''$ are calculated according to the formula:

$w(v') = \frac{w(v)}{2} + \frac{1}{2}|w(F_0\partial G) - w(G_0\partial F)|$

$w(v) = 1$

$w(G_0\partial F) = w(uzku) = uzk = 3$

$w(F_0\partial G) = 2 + 1 = 3$

$w(v') = \frac{1}{2} + \frac{1}{2}|3 - 3| = \frac{1}{2} + \frac{1}{2} \times 0 = \frac{1}{2} + 0 = \frac{1}{2} = 0.5$

$w(v'') = 1 - w(v') = 1 - 0.5 = 0.5$

Let us divide the weight of the second common variable $u$ with the best cost reduction.

$w(F_0) = w(xy) + w(pm) = 2$

$w(\partial F) = w(kv') = 1 + 0.5 = 1.5$

$w(G_0) = w(v''b) = 0.5 + 1 = 1.5$

$w(\partial G) = w(z) = 1$

$w(G_0\partial F) = w(v''bkv') = w(vbk) = 3$

$w(F_0\partial G) = 2 + 1 = 3$

$w(F_0G_0) = 2 + 1.5 = 3.5$

$w(u\partial F\partial G) = w(ukv'z) = 3.5$

$w(m_0) = \min\{w(F_0G_0), w(u\partial F\partial G)\} = \min\{3.5, 3.5\} = 3.5$

$2w(m_0) = 2 \cdot 3.5 = 7$

By using the best cost reduction we check if the following inequalities hold:

$$|w(G_0\partial F) - w(F_0\partial G)| < w(u) < 2m_0 - w(F_0\partial G) - w(G_0\partial F)$$

$$|3 - 3| < 1 < 7 - 3 - 3$$

$$0 < 1 < 1$$

According to the results of calculations the condition is not evaluated to true, therefore in division of the common variable's weight $u$ the defect $\triangle$ will equal 0.

Let us calculate the optimal values $u'$ and $u''$ with the best weight division:

$w(u') = \frac{w(u)}{2} + \frac{1}{2}|w(F_0\partial G) - w(G_0\partial F)|$

$w(u) = 1$

$w(G_0\partial F) = w(v''bkv') = vbk = 3$

$w(F_0\partial G) = 2 + 1 = 3$

$w(u') = \frac{1}{2} + \frac{1}{2}|3 - 3| = \frac{1}{2} + \frac{1}{2} \times 0 = \frac{1}{2} + 0 = \frac{1}{2} = 0.5$

$w(u'') = 1 - w(v') = 1 - 0.5 = 0.5$

Let us divide the cost of the second common variable $u$ with the worst cost reduction. With the worst cost reduction, there are 2 cases are examined where:

$$\text{1) } w(v') = 0, w(v'') = 1$$
$$\text{2) } w(v') = 1, w(v'') = 0$$

If it is known that the values are $w(v') = 0, w(v'') = 1$, then:
$w(F_0) = w(xy) + w(pm) = 2$
$w(\partial F) = w(kv') = 1 + 0 = 1$
$w(G_0) = w(v''b) = 1 + 1 = 2$
$w(\partial G) = w(z) = 1$
$w(G_0 \partial F) = w(v''bkv') = w(vbk) = 3$
$w(F_0 \partial G) = 2 + 1 = 3$
$w(F_0 G_0) = 2 + 2 = 4$
$w(u \partial F \partial G) = w(ukv'z) = 3$
$w(m_0) = \min\{w(F_0 G_0), w(u \partial F \partial G)\} = \min\{4, 3\} = 3$
By using the worst optimization we check if condition is completed:

$$\begin{cases} w(G_0 \partial F) < w(u) + w(F_0 \partial G), \\ w(G_0 \partial F) < m_0 = \min\{w(F_0 G_0), w(u \partial F \partial G)\} \end{cases}$$

or

$$\begin{cases} w(F_0 \partial G) < w(u) + w(G_0 \partial F), \\ w(F_0 \partial G) < m_0 = \min\{w(F_0 G_0), w(u \partial F \partial G)\} \end{cases}$$

$$\begin{cases} 3 < 1 + 3, \\ 3 < 3 \end{cases} \quad \begin{cases} 3 < 4, \\ 3 < 3 \end{cases}$$

or

$$\begin{cases} 3 < 1 + 3, \\ 3 < 3 \end{cases} \quad \begin{cases} 3 < 4, \\ 3 < 3 \end{cases}$$

None of conditions is completed.
If it is known that the values are $w(v') = 1, w(v'') = 0$, then:
$w(F_0) = w(xy) + w(pm) = 2$
$w(\partial F) = w(kv') = 1 + 1 = 2$
$w(G_0) = w(v''b) = 0 + 1 = 1$

$w(\partial G) = w(z) = 1$
$w(G_0 \partial F) = w(v''bkv') = w(vbk) = 3$
$w(F_0 \partial G) = 2 + 1 = 3$
$w(F_0 G_0) = 2 + 1 = 3$
$w(u \partial F \partial G) = w(ukv'z) = 4$
$w(m_0) = \text{min}\{w(F_0 G_0), w(u \partial F \partial G)\} = \text{min}\{3, 4\} = 3$

$$\begin{cases} w(G_0 \partial F) < w(u) + w(F_0 \partial G), \\ w(G_0 \partial F) < m_0 = \text{min}\{w(F_0 G_0), w(u \partial F \partial G)\} \end{cases}$$

or

$$\begin{cases} w(F_0 \partial G) < w(u) + w(G_0 \partial F), \\ w(F_0 \partial G) < m_0 = \text{min}\{w(F_0 G_0), w(u \partial F \partial G)\} \end{cases}$$

$$\begin{cases} 3 < 1 + 3, \\ 3 < 3 \end{cases} \quad \begin{cases} 3 < 4, \\ 3 < 3 \end{cases}$$

or

$$\begin{cases} 3 < 1 + 3, \\ 3 < 3 \end{cases} \quad \begin{cases} 3 < 4, \\ 3 < 3 \end{cases}$$

In this case none of conditions is also completed, therefore, in division of the common variable $u$ in the worst way, the defect $\triangle$ equals 0.

The last stage of the calculation is dedicated to defining the common defect. Taking into consideration the calculations it is now known that in the best cost reduction of the weight of common variables:
$w(u') = 0.5$, $w(u'') = 0.5$ and $\triangle_u = 0$
$w(v') = 0.5$, $w(v'') = 0.5$ and $\triangle_v = 0.5$
For calculating the common defect, it is necessary to find out the sum of the defect received from the division of the value of variable $v$ and the defect received in division of variable $u$.

$$\triangle_b = \triangle_u + \triangle_v = 0 + 0.5 = 0.5$$

It is also known that in the worst distribution of the value of common variables:
$\triangle_v \leq 1$ and $\triangle_u = 0$
For calculating the common defect, it is necessary to find out the sum of the defect received from the division of the value of variable $v$ and the defect received in division of variable $u$.

$$\triangle_w = \triangle_u + \triangle_v = 0 + 1 = 1$$

If we know the value of defects with the best and the worst cost reduction of the value division, we can calculate the difference between the best and the worst value. $\mid \triangle_b - \triangle_w \mid = |\frac{w(v)}{2} - w(v)| = \frac{w(v)}{2} = \frac{1}{2} = 0.5$, where $\triangle_b$ - defect with the best weight's division and $\triangle_w$ - defect with the worst weight's division.

The process of the best common values' distribution will be examined further. If we have necessary optimal values for $w(u'), w(u''), w(v'), w(v'')$, we can divide the common variables $u$ and $v$ one by one. At first stage one common variable is eliminated. When calculating the defect cost reduction method has been applied towards the variable $v$. It is known that the optimal values for the variable $v$ are: $w(v') = 0.5$ and $w(v'') = 0.5$
The attack tree is transformed as shown in Figure 4.6.

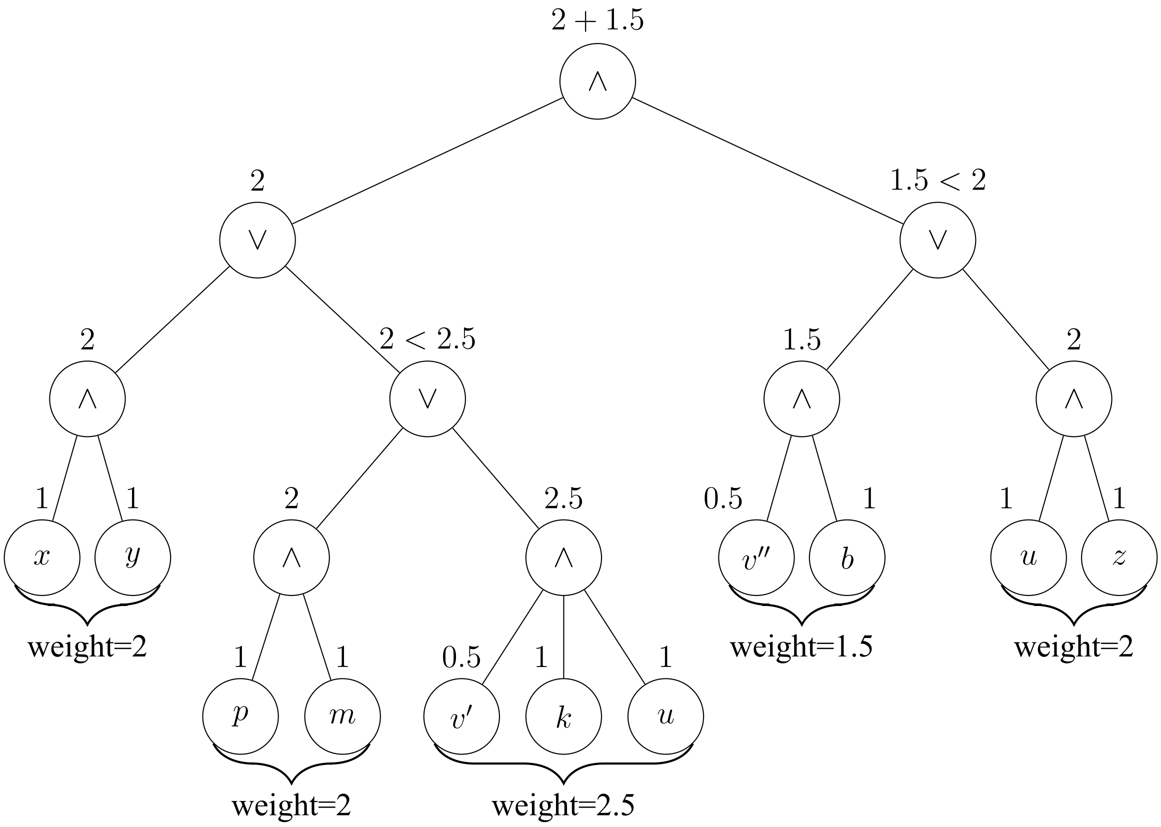

Figure 4.6: Attack tree after dividing the common variable $v$ by $v'$ and $v''$

After dividing the weight of common variable $v$, the solution for the function $\phi' = (F'G')$ is the value with the minimum weight from all possible solutions of the function:
$w(\phi) = \min\{w(xyv''b), w(xyuz), w(pmv''b), w(pmuz), w(v'kuv''b), w(v'kuz)\} =$
$= \min\{3.5, 4, 3.5, 4, 4, 3.5\} = 3.5$
Taking into consideration the received results, the cheapest combinations are the following ones: $(x \wedge y) \wedge (v'' \wedge b)$, $(p \wedge m) \wedge (v'' \wedge b)$ and $(v' \wedge k \wedge u) \wedge (u \wedge z)$ , weight $w$ equals 3.5, which is not an exact result in the end.

Then it is necessary to divide the variable $u$ by $u'$ and $u''$. It has been calculated earlier that the optimal meanings for $w(u')$ and $w(u'')$ are the values that equal $w(u') = 0.5$ and $w(u'') = 0.5$. After dividing the variable $u$, the attack tree looks like as it is shown in Figure 4.7.

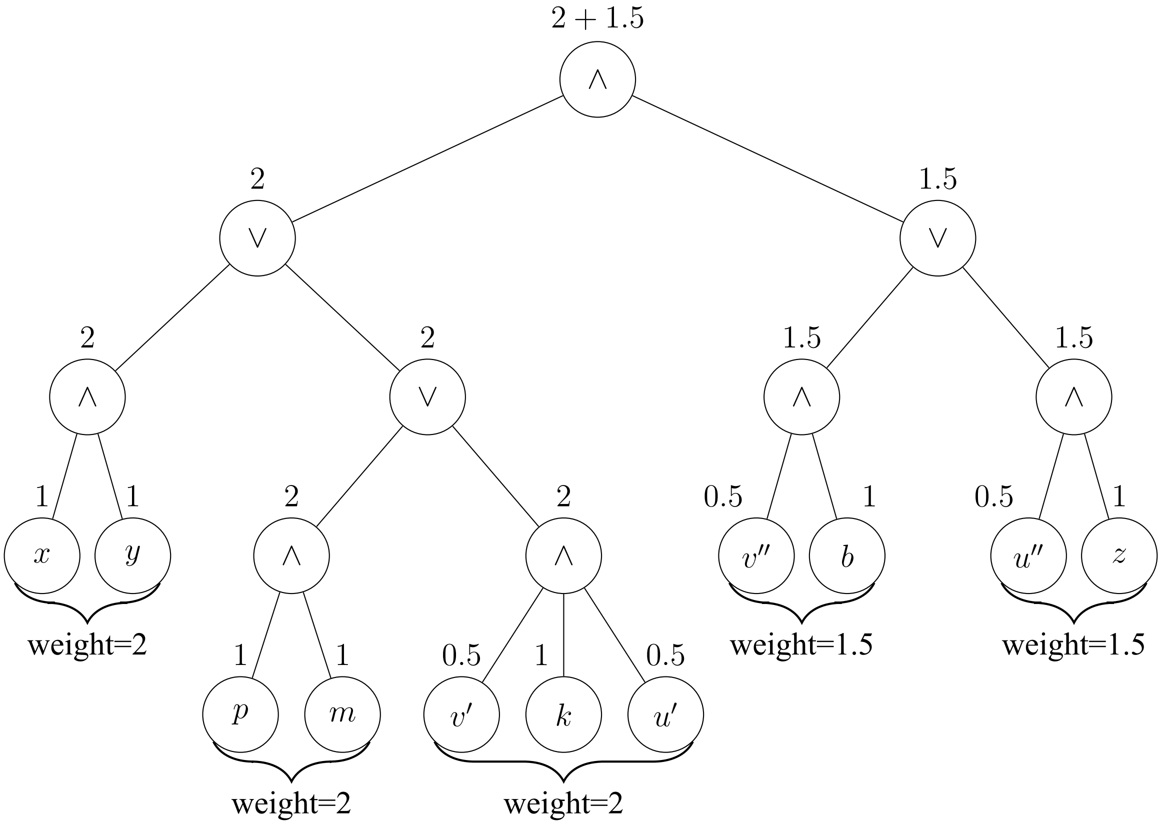

Figure 4.7: Attack tree after dividing the common variable $u$ by $u'$ and $u''$

After dividing the value of the common variable $u$, solution for the function $\phi' = (F'G')$ is the value with the weight that is minimal from all possible solutions of the function:
$w(\phi) = \min\{w(xyv''b), w(xyu''z), w(pmv''b), w(pmu''z), w(v'ku'v''b), w(v'ku'u''z)\} =$
$= \min\{3.5, 3.5, 4, 3.5, 3.5, 3.5\} = 3.5$
It also can be noticed that after dividing the variable $u$ the minimal results are $w(xyv''b) = 3.5$, $w(xyu''z) = 3.5$, $w(pmu''z) = 3.5$, $w(v'ku'v''b) = 3.5$ and $w(v'ku'u''z) = 3.5$, that corresponds to result from Theorem 1.

According to the data calculated, the conclusion can be made that after reduction the weight of 2 common variables the result differs from the exact result and has the defect $\triangle$.

## 4.6 Applying cost reduction to an attack tree

By using the method of propagation we can achieve the exact result in case of independent attack trees. Independent attack trees are the ones that do not contain common attacks. It is known that a defect appears in AND nods of the tree. The nods AND can be divided into two groups. The first group is linked with AND nodes containing the common attack. Cost reduction technique offers the division of common attacks' cost. Therefore, having divided common variable in AND node with common attacks, the node becomes a simple AND node without common attacks and can be related to the second group. Cost Reduction technique is an iterative process. During this process it is necessary to find all AND nodes with common attacks starting from the root node. After finding that AND node, the common variable is divided into two independent copies. In the left sub-tree the common variable is replaced for the first copy and in the right sub-tree the common variable is replaced for the second copy. Then, by going downwards from the root-node, the search of next AND-node with common attacks is being implemented. Division of the common variable happens in the same way described above. The procedure of AND-nodes with common attacks search and replacement of common variables is continued until the moment when AND-nodes with common attacks exist in the tree. When the tree gets rid of all AND-nodes with common attacks it becomes independent, which means that there are no more common attacks in the tree.

The process of logical elimination of common variables is shown further. There is a Boolean function $\phi = (vx + auz)(uy + bz)(qu + kmz)(qz + kum)$, where $F = (vx + auz)(uy + bz)$ and $G = (qu + kmz)(qz + kum)$. Attack tree before implementing the method of propagation looks as ashown in Figure 4.8. Top AND node contains 2 common variables - $u$ and $z$ (they are marked with red colour in Figure 4.8).

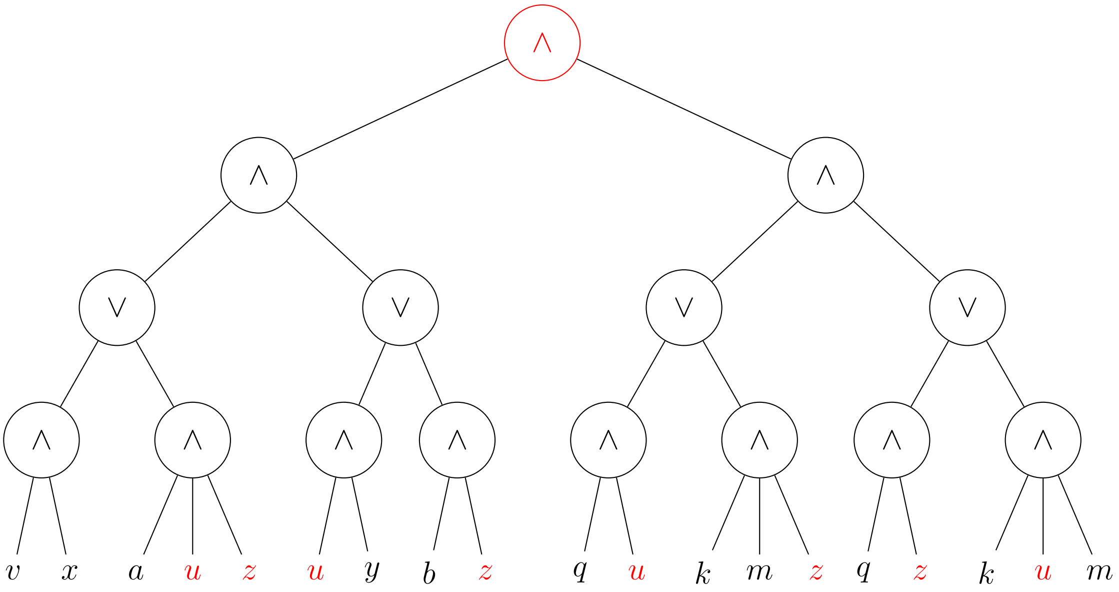

Figure 4.8: Attack tree before using propagation method

**Transformation 1.**

After substitution of $u \mapsto u'$ and $z \mapsto z'$ in the left sub-tree $F$, and $u \mapsto u''$ and $z \mapsto z''$ in the right sub-tree $G$, the attack tree gets the look shown in Figure 4.9 (changes are marked with red colour).

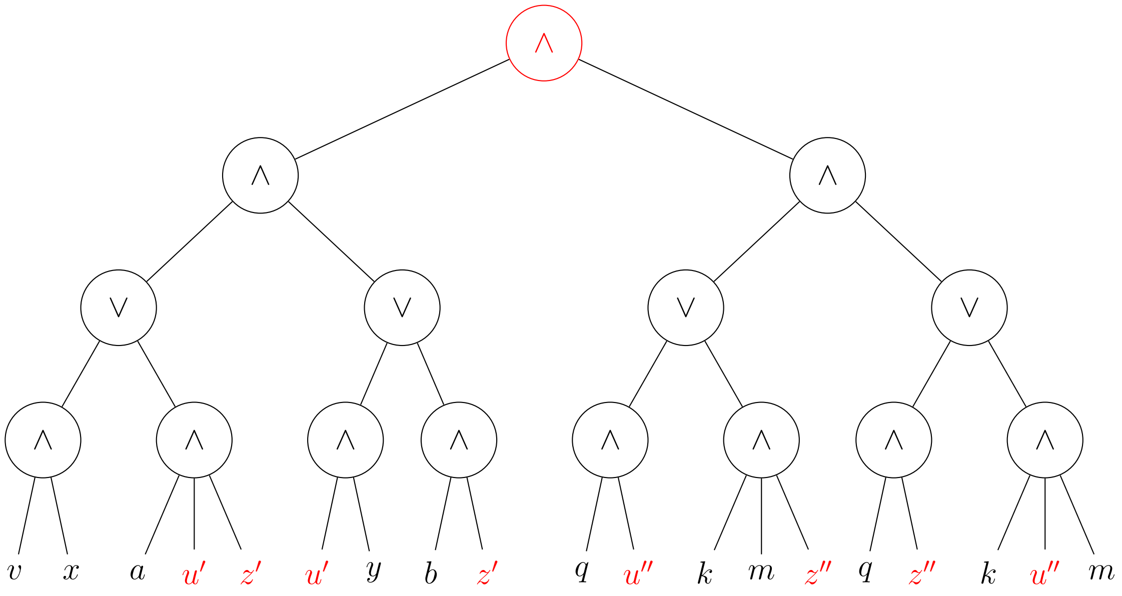

Figure 4.9: Attack tree after the 1-st iteration

After reducing the cost of common attacks, the top AND node does not contain any common attacks (dependent variables).

**Transformation 2.**

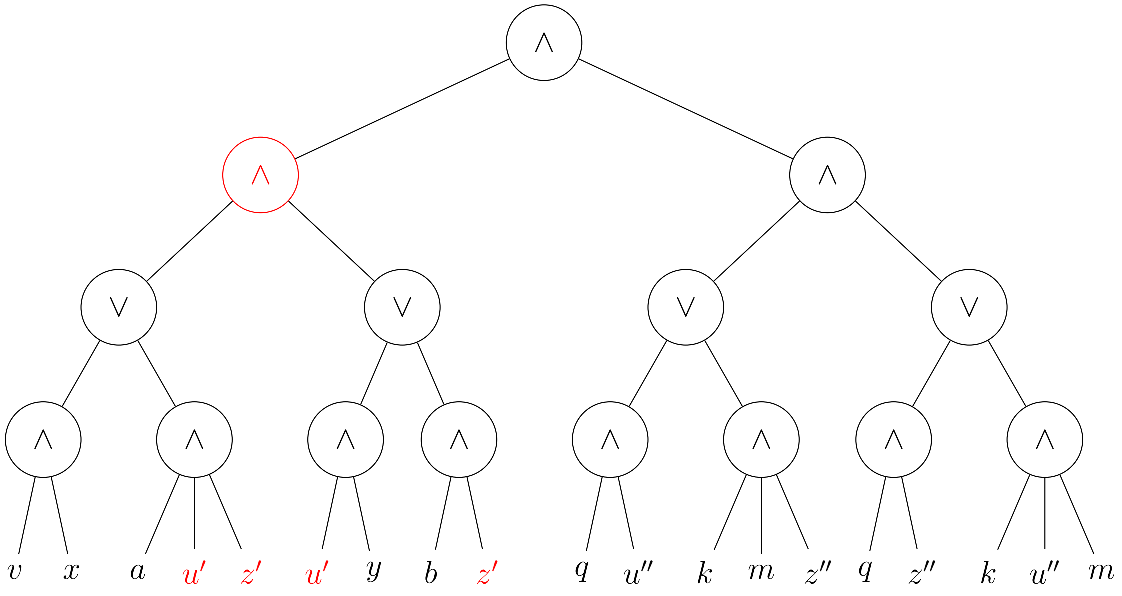

Figure 4.10: $F$ sub-tree

Further it can be seen that there are common variables $u'$ and $z'$ in the left AND root-node $F$ (Figure 4.10).

We conduct the substitution of $u' \mapsto u'{}'$ and $z' \mapsto z'{}'$ in the left sub-tree $F$, and substitution of $u' \mapsto u'{}''$ and $z' \mapsto z'{}''$ in the right sub-tree $F$. After implementing the method of cost reduction we get the attack tree shown in Figure 4.11.

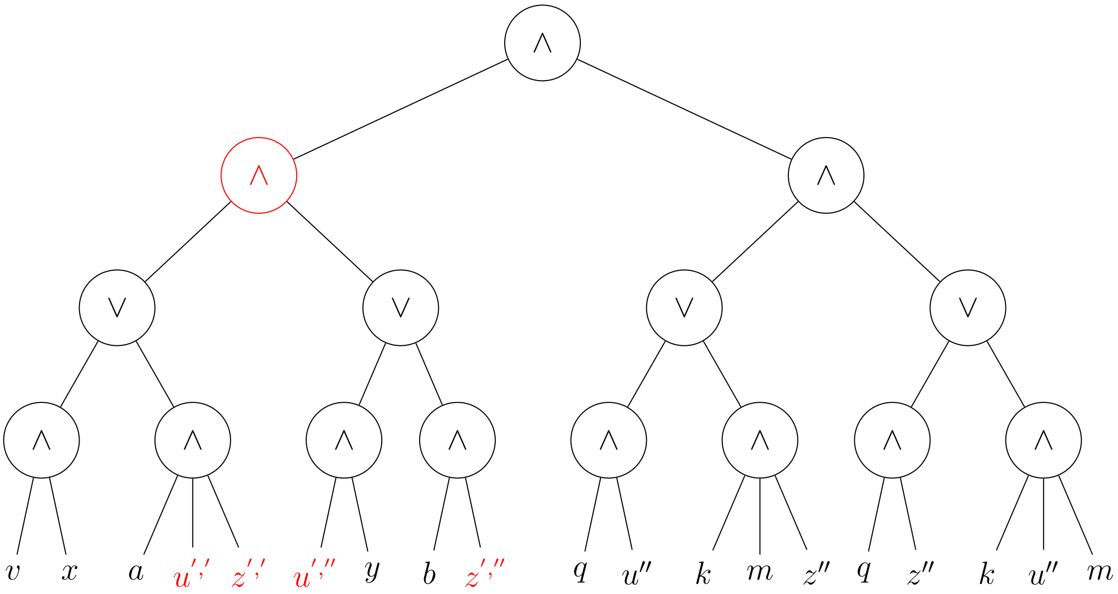

Figure 4.11: Attack tree after the 2-nd interation

After implementing the second iteration there are no more common attacks (dependent variables) in the root-node AND of the left sub-tree $F$.

**Transformation 3.**

Let us look at the right sub-tree $G$, with the root-node AND. Sub-tree $G$ contains common attacks $q$, $u''$, $k$, $m$ and $z''$ (is marked with red colour in Figure 4.12).

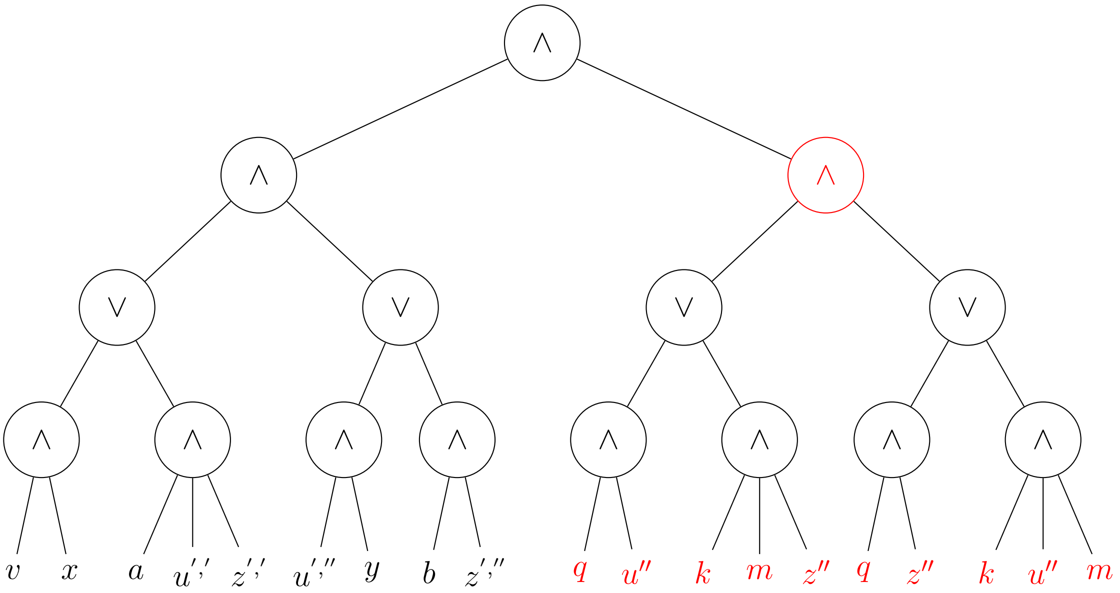
∧
∧ ∧
∨ ∨ ∨ ∨
∧ ∧ ∧ ∧ ∧ ∧ ∧ ∧
$v$ $x$ $a$ $u^{','}$ $z^{','}$ $u^{','',}$ $y$ $b$ $z^{',''}$ $q$ $u''$ $k$ $m$ $z''$ $q$ $z''$ $k$ $u''$ $m$

Figure 4.12: $G$ sub-tree

For the left sub-tree $G$ we conduct the following replacement: $q \mapsto q'$, $u'' \mapsto u^{'','}$, $k \mapsto k'$, $m \mapsto m'$ and $z'' \mapsto z^{'','}$. for the right sub-tree $G$ we do the replacement $q \mapsto q''$, $z'' \mapsto z^{'',''}$, $k \mapsto k''$, $u'' \mapsto u^{'',''}$ and $m \mapsto m''$. After implementing the method of cost reduction we get the attack tree shown in Figure 4.13.

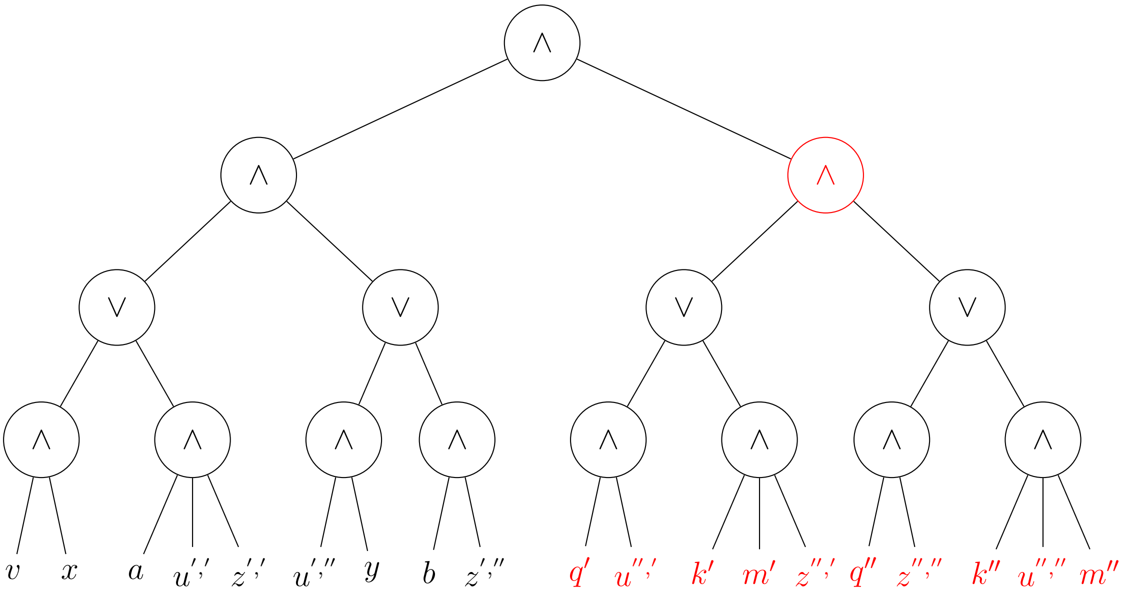
∧
∧ ∧
∨ ∨ ∨ ∨
∧ ∧ ∧ ∧ ∧ ∧ ∧ ∧
$v$ $x$ $a$ $u^{','}$ $z^{','}$ $u^{',''}$ $y$ $b$ $z^{',''}$ $q'$ $u^{'','}$ $k'$ $m'$ $z^{'','}$ $q''$ $z^{'',''}$ $k''$ $u^{'',''}$ $m''$

Figure 4.13: Attack tree after the 3-rd iteration

Having completed all iterations, we have received the attack tree without common attacks. We can apply propagation method and calculate the result in this tree.

# 5 Conclusion

The aim of the research was to find out if there is such division of the costs of 2 or more common attacks where cost reduction will provide the exact result. In the general case such division is impossible, as the counter-example showed that there is the best division of attacks' costs with existence of defect. This research has evaluated the accuracy of the cost reduction method, conditions under which the reduction defect exists, as well as the size of the defect. In the process of this research, the best and the worst optimization of the costs' division of attacks have been studied. It depends on the level of accuracy of the result, so we apply either occasional cost division or we optimize the division of common attacks' cost.

Therefore, if the difference between the best and the worst division is not big, we can, for instance, use symmetric division of the common variables' weights. By applying symmetric cost reduction we achieve the result is a short period of time, but the accuracy of the calculation may be insufficient. It also should be considered that the division of the attack costs in an arbitrary way can be wisely used if the number of common attacks is not big and their price is not high. If the number of common attacks is big or the costs are high, the size of the defect, when dividing the common costs in occasional way, can be high which is not acceptable. In this case it makes sense to invest a bit more time and optimize the attacks' cost division in a way when the value of divided variables become the best ones.

## 5.1 Open Questions and Future Work

We can calculate the result of analysis by using cost reduction technique in various ways. The first way is to optimize each common variable step by step, as shown in section 4.6, dividing not more than 1 variable at a time. Another way is to optimize all common variables as a multi-parameter optimization problem. Further it is necessary to investigate which way is optimal.

Apart from that, using step by step reduction of variables, there is still an open question if the result depends on the chosen ordering, in which costs of common variables are reduced. If the result is independent from the reduction ordering, the result will be the same, whatever ordering we use.

Another point for future research is to find out if the result is going to be the same in the case of semantically different attack trees that represent the same Boolean function. Different possible structures are shown in Figures 5.1 and 5.2.

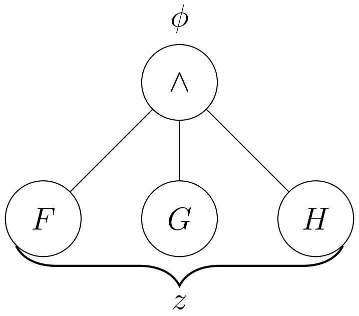

Figure 5.1: Function $\phi = FGH$, where $F, G, H$ depends on $z$

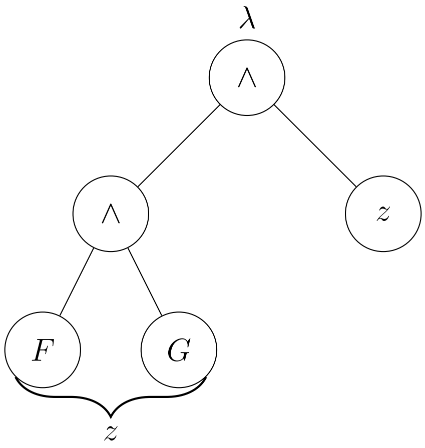

Figure 5.2: Function $\phi = (FG)H$, where $F, G, H$ depends on $z$

It is connected with the fact that the same task can modelled in a different way. If two Boolean functions are equivalent, the reliable method must yield the same utility no matter what structure of the tree is. If this is not the case, the cost reduction method is not reliable method.

# Bibliography

[1] Popper, K.R.: The Logic of Scientific Discovery (Routledge Classics). Routledge (2002, first published in 1934)

[2] Lenin, A.: Reliable and Efficient Determination of the Likelihood of Rational Attacks., TUT Press (2015) 21–32, 55–73, 75–94, 107–127

[3] Buldas, A., Laud, P., Priisalu, J., Saarepera, M., Willemson, J.: Rational choice of security measures via multi-parameter attack trees. In López, J., ed.: CRITIS. Volume 4347 of Lecture Notes in Computer Science., Springer (2006) 235–248

[4] Ionita, D.: Current established risk assessment methodologies and tools, University of Twente (2013)

[5] Miede, A., Nedyalkov, N., Gottron, C., König, A., Repp, N., Steinmetz, R.: A generic metamodel for IT security. In: ARES 2010, Fifth International Conference on Availability, Reliability and Security, 15-18 February 2010, Krakow, Poland. (2010) 430–437

[6] Kordy, B., Mauw, S., Radomirovic, S., Schweitzer, P.: Attack-defense trees. J. Log. Comput. **24**(1) (2014) 55–87

[7] Schneier, B.: Attack trees: Modeling security threats. Dr. Dobb's Journal **24**(12) (1999) 21–29

[8] Vesely, W.E., Goldberg, F.F., Roberts, N.H., Haasl, D.F.: Fault tree handbook. Technical report NUREG–0492. (1981)

[9] Weiss, J.D.: A system security engineering process. Volume 249., Proceedings of the 14th National Computer Security Conference (1991) 572–581

[10] Mauw, S., Oostdijk, M.: Foundations of attack trees. In Won, D., Kim, S., eds.: ICISC. Volume 3935 of Lecture Notes in Computer Science., Springer (2005) 186–198

[11] Geer, D., Hoo, K.S., Jaquith, A.: Information security: Why the future belongs to the quants. IEEE Security & Privacy **1**(4) (2003) 24–32

[12] Jürgenson, A., Willemson, J.: Computing exact outcomes of multi-parameter attack trees. In Meersman, R., Tari, Z., eds.: OTM Conferences (2). Volume 5332 of Lecture Notes in Computer Science., Springer (2008) 1036–1051

[13] Jürgenson, A., Willemson, J.: Serial model for attack tree computations. In: Information, Security and Cryptology - ICISC 2009, 12th International Conference, Seoul, Korea, December 2-4, 2009, Revised Selected Papers. (2009) 118–128

[14] Buldas, A., Stepanenko, R.: Upper bounds for adversaries' utility in attack trees. In Grossklags, J., Walrand, J.C., eds.: GameSec. Volume 7638 of Lecture Notes in Computer Science., Springer (2012) 98–117

[15] Buldas, A., Lenin, A.: New Efficient Utility Upper Bounds for the Fully Adaptive Model of Attack Trees. In Das, S.K., Nita-Rotaru, C., Kantarcioglu, M., eds.: GameSec. Volume 8252 of Lecture Notes in Computer Science., Springer (2013) 192–205