Safety optimization: a combination of fault tree analysis and optimization techniques

Frank Ortmeier, Wolfgang Reif

Angaben zur Veröffentlichung / Publication details:

Ortmeier, Frank, and Wolfgang Reif. 2004. “Safety optimization: a combination of fault tree analysis and optimization techniques.” Augsburg: Universität Augsburg.

Safety Optimization: A combination of fault tree analysis and optimization techniques

Frank Ortmeier Wolfgang Reif

***Abstract*— We present a new form of quantitative safety analysis - safety optimization. This method is a combination of fault tree analysis(FTA) and mathematical optimization techniques. With the use of the results of FTA, statistics, and a quantification of the costs of hazards, it allows to find the optimal configuration of a given system with respect to opposed safety requirements. Furthermore, the system may not only be examined for safety, but usability as well.**

We illustrate this method on a real-world case study: the height control system of the Elbtunnel in Hamburg. Safety optimization showed some significant problems in trustworthiness of the system, yielded optimal values for configuration of free parameters and showed possible modifications to improve the system.

***Index Terms*—fault tree analysis, dependability, optimization, safety analysis, embedded systems**

I. Introduction

In this paper we describe how mathematical optimization and safety analysis can be combined. We call this technique safety optimization. We applied this method on a case study - the height control system of the Elbtunnel in Hamburg and will illustrate the method on this example. This work has been developed within the ForMoSA project which is part of the german research foundations priority progam “Integrating software specification techniques for engineering applications”.

Many modern books about safety [6][7][14] do not cover quantitative methods or only cover them on the fringes. This is, because quantitative methods usually rely heavily on statistics. So they are often seen as a problem of mathematics. However, it can improve safety analysis a lot, if good interfaces between mathematics and statistics are provided. Safety optimization is an enhancement to the well-known fault tree analysis (FTA) [16][2] which makes it easy to integrate statistics and optimization techniques into safety analysis. The new aspect - compared to traditional FTA - is that we build a statistical model of the environment in addition to the components' failure probabilities known from quantitative FTA. With this model, the results of quantitative FTA, and a cost function optimal configuration values for free parameters may be found as solutions of an optimization problem. Many real world applications have free parameters, which influence safety requirements: the tolerance of a speed indicator, accepted time delay between request and answers or the average maintenance interval are all free parameters of different systems. Such parameters are normally chosen on a basis of previous experience and fine tuned once the system starts working. However, bad choices only become obvious when some hazards occur. So it would be very helpful, if these parameters could be estimated in advance. This is what safety optimization does.

In Sect. II we give a brief introduction of FTA including quantitative FTA. The theoretical foundations of safety optimization are part of Sect. III. An application of safety optimization on a real world case study is presented in Sect. IV. Limitations and future work are presented in Sect. V. Section VI concludes the paper.

II. Fault Tree Analysis

In this section we start with a brief introduction into fault tree analysis(FTA). FTA is a top down technique to determine the possible basic component failures (primary failures) of a bad or catastrophic situation which must be avoided. This situation is called hazard.

The hazard or top event is always the root of the fault tree and primary failures are its leaves. All inner nodes of the tree are called intermediate events. Starting with the top event the tree is generated by determining the immediate causes that lead to the top event. They are connected to their consequence through a gate. The gate indicates if all (and-gate) or any (or-gate) of the causes are necessary to make the consequence happen. the INHIBT-gate states, that the cause is only critical if some environmental condition holds. Unlike all other nodes of the fault tree, this condition must not be a failure or undesired event. The leaves of a fault tree are primary failures which are not investigated further. Figure 1 shows the symbols for fault tree gates.

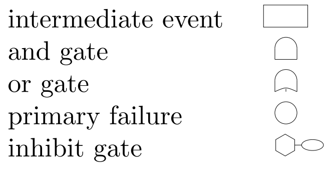

Fig. 1. fault tree symbols

The fault tree in figure 2 is part of the FTA of the Elbtunnel height control system which we will describe in detail in Sect. IV. This tree describes, that the immediate causes of the top event - collision - are that either the driver ignores some stop signals OR (this means the causes are connected through an OR-gate) that the signals are not turned on. The first cause is a primary failure. We can do nothing about it, but to disbar the driver from his license. The second cause is an intermediate event. Its immediate causes are a) that the signal lights are out of order or b) the signals were not activated. Again the first one is a

Lehrstuhl für Softwaretechnik und Programmiersprachen — Universität Augsburg — D-86135 Augsburg
email: {ortmeier,reif}@informatik.uni-augsburg.de

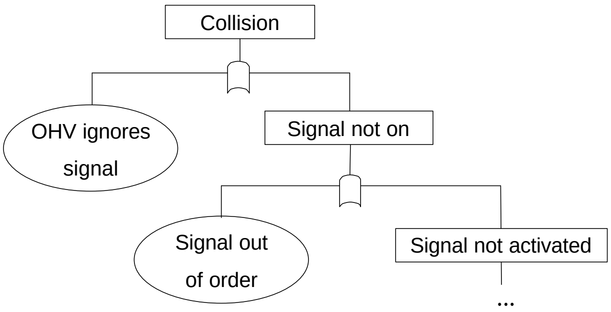

Fig. 2. Collision fault tree

primary failure and the second is an intermediate event, which has to be investigated further. This procedure has to be applied recursively to all causes until the desired level of granularity is reached (this means all causes are primary failures that won't be investigated further).

In the remainder of this section we will describe the most important terms of FTA in such detail, that the concept of safety optimization can be explained later.

A. Primary Failures and Hazards

For quantitative analysis roots and leaves—i.e. hazards and primary failures—are important. Let $F := \{PF_1, .., PF_n\}$ be the set of all primary failures PF_i and $H := \{H_1, .., H_m\}$ the set of all hazards H_i under consideration.

For each hazard H_i a separate fault tree has to be constructed, that describes which combination of basic causes (= primary failures) may be the reason for the hazard.

B. Minimal cut sets

An interesting result of fault tree analysis are the minimal cut sets. A cut set $CS_{H_i} \subseteq F$ for a hazard H_i is a set of primary failures, which together form a threat. This means if all primary failures of the cut set take place, then the hazard may occur.

Minimal cut sets $MCS_{H_i} \subseteq F$ for a hazard H_i are cut sets, such that preventing one element of each minimal cut set prevents the hazard from occurring. The set of all minimal cut sets for a hazard H_i may be automatically generated from the fault tree[16]. We call this set $MCSS_{H_i}$. So minimal cut sets describe qualitatively the dependency between hazards and primary failures. Minimal cut sets may be derived from the structure of the fault tree automatically.

Even more interesting for real world applications is the quantitative one i.e. the dependency between the probability of occurrence of the cut sets and the hazard. This question can be answered with quantitative FTA.

C. Calculating probabilities

For calculating probabilities we use a standard formula for calculating hazard probabilities from fault trees[16]. It calculates the probability of a cut set as the product of the probabilities of all its elements. The hazard's probability is calculated as the sum of all its minimal cut sets probabilities. So the probability of a hazard H_i is calculated as:

$$P(H_i) := \sum_{MCS \in MCSS_{H_i}} P(MCS) \tag{1}$$

$$where\ P(MCS) := \prod_{PF \in MCS} P(PF)$$

This formula is widely used in engineering and broadly accepted, but uses some assumptions about statistical independence. All primary failure are assumed to be pairwise independent. This hold for many applications. If statistical correlation has to be examined, FTA is not a good choice and another approach like common cause analysis or—on the formal side—stochastic model checking [1] has to be used and the probability of the minimal cut sets and hazards have to be calculated separately. this formula alse neglects second and higher-order terms in the sum. This is in practice no problem as failure probabilities are very small.

In this paper we stick with the assumption of statistical independence and use the standard formulae as starting point for our extension of FTA.

D. Generalizations

So far this is the standard method of applying quantitative FTA. But for our purposes this is not enough. As (i) for this point of view the worst case is always assumed. This means, all other environmental inputs are as "bad" as possible. Another deficiency is (ii) the use of fixed probabilities for failures. In reality these probabilities are usually not constant, but rather depend on some parameters. To overcome these problems, we generalize quantitative FTA by introducing two new types of probabilities: constraint probabilities and parameterized failure probabilities.

D.1 Constraint probabilities

Most of the cut sets cause the hazard only if one or more constraints are fulfilled. Sometimes a constraint must hold for all cut sets and sometimes only for some of them. For example the failure of a critical cooling unit is only dangerous if the system which has to be cooled is working. If it is turned off, then the failure of such a cooling unit will not have any effect onto the super system. There exist quite some extension for FTA which consider such effects on a qualitative basis. The qualitative dependence between such constraints, cut sets and hazards is then integrated in the fault tree with so called INHIBIT-gates (see figure 1). An INHIBIT-gate states a condition or constraint which has to be fulfilled such that the failure cause makes the consequence happen. This condition need not necessarily be a failure, but can also be some environmental influence. Although some types of FTA respect such dependencies in a qualitative manner, they are in general neglected for quantitative FTA.

We introduce constraint probabilities for quantitative analysis which reflect how probable it is that the inputs from the environment are "bad" enough to make the hazard happen. So we refine definition of a cut sets probability to get a better approximation:

$$P(CS) := P(Constraints) \prod_{PF \in \mathrm{CS}} P(PF) \qquad (2)$$

If one chooses P(Constraints)=1, it means the environment always behaves as bad as possible and one gets the same formula as before. However, if one can estimate P(Constraints) a priori, then the results will be much more precise. This estimation can be approximated by calculating the probabilities of all conditions in INHIBIT-gates along the paths through the tree from the hazard to the elements of the cut sets. An upper bound fpr the constraint probability is then the product of all conditions' probabilities if statistical independence holds; if not then the maximum is an upper bound for it.

In practice these numbers are really hard to calculate exactly. So most of the time they are only approximated. But even if they can not be approximated very well, they still may be a great help for safety analysis. This is, because variation of the constraints allows to examine the behavior of the system in different working environments. This can give good advice, when trying to estimate how the system will scale in future. The benefit of this last methodology becomes obvious in the case study of Sect. IV.

D.2 Parameterized probabilities

The second important generalization we made is that we not only use constant failure probabilities for primary failures, but allow parameterized probabilities. This means, if the probability of primary failures PF (e.g. a relay fails to close) depends on some parameter X (e.g. the spring tension of the relay), we use a functional mapping between X and P(PF) and write P(PF)(X).

$$P(PF):\ Domain(X) \rightarrow [0,1]$$

In principle, there is no restriction on the domain of X, as they only affect which methods are applicable for the solution of the resulting optimization problem. But finite and discrete domains are in general less interesting and rare. In practice P(PF) is usually a (continuous) probabilistic distribution. If the probabilities depend on more then one parameter, then take X as the vector of all involved parameters.

All instances of failure probabilities are substituted with the according function. The algorithm for calculating the probability of a cut set is not changed. So the probability of a cut set is then also function of one or more variables. The same is true for the hazards' probabilities and formula 1 now rewrites to:

$$\begin{aligned} P(H_i)(X) &= \sum_{MCS \in MCSS_{H_i}} P(MCS)(X) & (3) \\ P(H_i)(X) &= \sum_{MCS \in MCSS_{H_i}} \prod_{PF \in MCS} P(PF)(X) & (4) \end{aligned}$$

Note, that the probabilities of cut sets and hazards are no longer fixed numbers, but rather functions of the free parameters of the system. We call these functions parameterized probabilities.

III. Safety Optimization

In this section we describe the combination of quantitative fault tree analysis and optimization techniques, which leads to safety optimization. The basic idea is as simple as effective. In practice for most systems safety is a tradeoff between different undesired events. For example in aviations the main goal of a pre-flight safety check on an airplane before start is to make sure the aircraft is working correctly and will not crash. However, another important goal of the safety check is that an aircraft, which allows safe flight must not fail the check. Assume that one part of the check is aberration of the air speed indicator. Then it is obvious that the smaller the allowed tolerance is, the safer the airplane operation will be. On the other hand to small acceptable tolerances will result in many safe aircraft failing the pre-flight check and thus in delay or canceled flights. So what is the solution? Its of course some middle value between zero tolerance and arbitrary tolerance[1].

This is exactly the point where safety optimization works. It uses mathematical optimization to find the best value for this tolerance parameter. To do this we only need one more information: the cost function.

A. *Cost function*

To do mathematical analysis, a cost function is needed. A cost function describes the total costs that all hazards together cause in average to the operator. This is done by risk assessment. The cost of each hazard will be defined. It is common practice—even as it may seem un-ethical—to do this in cash (e.g. one dead person is calculated with 2.7M \$ in US railways organizations).

More important for the operators of the system are the mean costs. These are the costs, which have to be expected. These costs depend on the probability of occurrence and absolute cost of the hazard. In many cases cost functions are simply the weighted sum of hazards' probabilities. Where the weights represent the costs associated with each hazard.

$$f_{cost}(P(H_1),..,P(H_m)) := \Sigma_{i=1}^{m} Cost_{H_i} P(H_i) \qquad (5)$$

In our approach, the probabilities of the hazards $P(H_i)$ are not, as for standard quantitative FTA, necessarily constants, but rather functions of free parameters $X_1...X_l$[2] if the contributing cut sets are described by parameterized probabilities (see 3). So the cost function can be expressed as function of the free parameters $X_1...X_l$. This allows us

[1] We are not arguing for safety leaks which originated in design flaws, to be left open because of the costs. This approach only addresses hardware failure which can not ultimately be avoided.

[2] In reality, not every hazard H_i depends on all free parameters $X_1...X_l$, but rather only on a subset. Therefore, we write $P(H_i(X_{i,1},...,X_{i,n_i}))$ in equation 6.

to use optimization techniques.

$$f_{cost}(X_1,...,X_l) := \Sigma_{i=1}^{m} Cost_{H_i}(P(H_i(X_{i,1},...,X_{i,n_i})) \quad (6)$$

B. Mathematical optimization

Once the cost function is defined the problem is not any more a safety analysis problem but only a mathematical one. The goal is to choose the free parameters $X_1...X_l$ such that the cost function is minimized. So the problem is:

$$Find\ (x_1,...x_l)\ such\ that:\ f_{cost}(x_1,...,x_l) = min_{X_1,...,X_l} f_{cost}$$

To guarantee the existence of the minimum we restrict the real value domains to be compact intervals. This problem can then be solved with different methods. In simple cases analytical solutions may be found. If the problem is more complex and the cost function still smooth enough (e.g. twice-continuously differentiable) then there exist a lot of algorithms from the domain of nonlinear programming to solve the problem. The most simple one is the gradient method which finds local minima by calculating gradients iteratively and always following the steepest descent. But there exists a wide variety of more elaborate and efficient algorithms. A good introduction to optimization of nonlinear problems may be found in [4] and [8].

If there are only two free variables (as in the following example) and the functions are smooth, than the solutions may be found by using a 3D plot of the cost function and zooming into it. Even if a specific optimization problem is neither analytically nor numerically solvable, this method can yield some results by testing possible combinations. It is possible to test large number of combinations in very short time. So this technique gives a good impression about the quantitative dependencies between mean costs and free parameters.

IV. An example: The Elbtunnel

In this section we will apply the presented technique step by step to a real world case study: the height control system of the Elbtunnel in Hamburg. We will give a short summary of the problem and show the most interesting parts of the analysis and the results.

The Elbtunnel is a road tunnel beneath the river Elbe in Hamburg. Until end 2002 this tunnel consisted of three identical tubes with two driving lanes each. Late 2002 a new, fourth tube went into operation. This new tube is different. It is bigger and allows the crossing of vehicles, which are too big for the old three tubes. Furthermore, the tunnel has a dynamic traffic control which switches driving directions from north to south and vice versa in the two middle tubes according to traffic needs. This is done by an electronic traffic control, underground signals, and electric road signs. Figure 3 shows the layout of the tunnel. The new, bigger tube is tube number 4.

We will consider only a small part of the whole project, the height control. This system must ensure, that vehicles, which are to high for the old three tubes, may only enter the new tube. Otherwise an emergency halt should be signaled automatically.

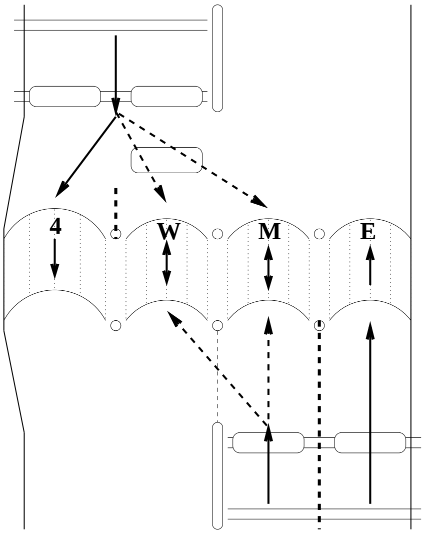

Fig. 3. Layout of the tunnel

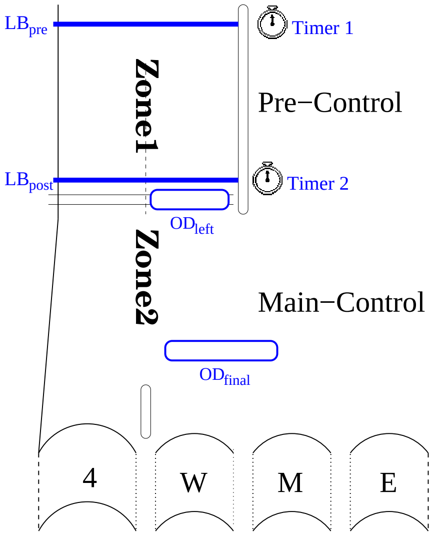

Fig. 4. Control of the northern entrance

A. The height control system

In the following, we will distinguish between normal cars, *high vehicles* (HVs) e.g. normal trucks or buses, which may drive through all tubes and *overhigh vehicles* (OHVs) e.g. extra large trucks, which can only drive through the new, fourth tube.

In this paper we will only discuss the height control system for the northern entrance of the tunnel; the southern end is less complex, as no OHV at all may pass the tunnel in that direction (they have to take a nearby bridge). The control for the northern entrance is shown in figure 4.

The system uses two different types of sensors. Light barriers (LB) are scanning all lanes of one direction to detect, if an OHV passes. For technical reasons they cannot be installed in such a way, that they only supervise one lane. Therefore overhead detectors (OD) are necessary to detect, on which lane an OHV passes. The ODs can distinguish vehicles (e.g. cars) from high vehicles (e.g. buses, trucks), but not HVs from OHVs (but light barriers can!). If the height control detects an OHV heading towards a different than the fourth tube, then an emergency stop, locking the tunnel entrance, is signaled.

The idea of the height control is, that the detection starts, if an OHV enters zone1 (see figure 4) at light barrier $\mathrm{LB}_{\mathrm{pre}}$. This results in activation of $\mathrm{LB}_{\mathrm{post}}$. To prevent unnecessary alarms through faulty triggering of $\mathrm{LB}_{\mathrm{pre}}$, $\mathrm{LB}_{\mathrm{post}}$ will be switched off after expiration of a 30-minutes-timer (called timer1). Road traffic regulations require, that after $\mathrm{LB}_{\mathrm{pre}}$ both HVs and OHVs have to drive on the right lane through tunnel 4. If nevertheless an OHV drives on the left lane towards the west-tube, detected trough the combination of $\mathrm{LB}_{\mathrm{post}}$ and $\mathrm{OD}_{\mathrm{left}}$, an emergency stop is triggered.

If the OHV drives on the right lane through $\mathrm{LB}_{\mathrm{post}}$, it is still possible for the driver to switch in zone2 to the left lanes and drive to the west- or mid-tube. To detect this situation, the height control uses the $\mathrm{OD}_{\mathrm{final}}$ detector. This sensor is activated by $\mathrm{LB}_{\mathrm{post}}$ and will be kept active for the next 30 minutes (called timer2).

For safe operation it is necessary, that after the location of $\mathrm{OD}_{\mathrm{final}}$ it is impossible to switch lanes. We already presented a safety analysis of this system in [10] consisting of verification of functional properties and FTA. With formal verification using the SMV-tool[9] we discovered a design flaw, which resulted in a possible hazard if two OHVs passed LB_{pre} simultaneously. After presenting solutions to this problem, we could proof functional correctness for the collision hazards. For false alarms we could show, that it there were only two possibilities for false alarms: a high vehicle at OD_{Left} while an OHV passes OD_{Right} or a HV at OD_{Final}. After consultation with the engineers both problems were said to be of minor importance, because in both cases a driver neglects road traffic regulations.

We also discovered by FTA that the runtimes of the timers are crucial to the system, but could not assess their exact effects onto the system. We will now only focus on finding optimal values for the runtimes of these timers and evaluate the design in different environments.

B. Statistical model of the Elbtunnel

B.1 Primary failure and hazards

There are two different, interesting hazards for the Elbtunnel height control which we analyzed - the collision (H_{Col}) of an OHV with the tunnel entrance and the tripping of a false alarm (H_{Alr}). For both a separate fault tree has to be built. It is clear that it is not possible to minimize both risks at the same time. We could also give formal proof for this by doing formal FTA [12]. The primary failures can be divided into four different types:

- False Detection (FD): The sensor *does* indicate a vehicle, although there is *none.* Possible for all sensors.
- Miss Detection (MD): The sensor *does not* indicate a vehicle, although there is *one.* Only possible for microwave sensors.
- Overtime (OT): Actual driving time of an OHV exceeds the runtime of a timer. Possible for timer 1 and timer 2.
- High vehicles (HV): A high vehicle beneath an overhead detector is interpreted as an OHV.

FDs are possible for all sensors, while MD and HV is only interesting for OD-type sensors. We write $FD_{OD_{final}}$ as abbreviation for false detection at overhead detector OD_{final} and analogously for all other sensors. Overtime failures can occur in zone 1 and zone 2. We write OT_1 resp. OT_2.

Note, that the last item (HV) is not a failure in the traditional sense, as overhead detectors can not distinguish between high vehicles and OHVs, high vehicles at the location of the sensors are (incorrectly) interpreted as OHVs. For the control system this has the same effect as a FD of the sensor.

In notation of Sect. II and III we get: $H = \{H_{Col}, H_{Alr}\}$ and F=$\{HV_{OD_{left}}, FD_{OD_{left}}, MD_{OD_{left}}, HV_{OD_{final}}, FD_{OD_{final}}, MD_{OD_{final}}, OT_1, OT_2, FD_{LB_{pre}}, FD_{LB_{post}}\}$.

We do not consider failures outside the detection system like broken stop signals or drivers who are ignoring an emergency halt. Defects in the micro-controller and the timers have also been neglected. It is estimated, that failure of these components is by orders of magnitude (at least 2) smaller then misdetection or false detection of the sensory devices.

B.2 Minimal cut sets

For the hazard “collision” almost all cut sets are single point of failures. The two most important ones are those, that are caused by traffic jams in zone 1 resp. zone 2. In our notation: $\{OT_1\}$ and $\{OT_2\}$. The other minimal cut sets are left out here for better readability. We write P_{const_1} for their combined probability.

False alarms may be triggered by $\{HV_{OD_{left}}\}$, $\{HV_{OD_{final}}\}$, $\{FD_{OD_{left}}\}$, or $\{FD_{OD_{final}}\}$. These are also all single point of failures. It is interesting to note, that all these failures are only then single point of failures, if there is an OHV present in the controlled area. We will only consider $\{HV_{OD_{final}}\}$ in detail in this paper. In fact it turns out that this will be the dominating factor in the hazards H_{Alr} overall probability by two orders of magnitude. We accumulate the probability of all other cut sets in P_{const_2}. So we get the following formulae for the hazards probabilities (see Sect. II-C):

$$\begin{aligned} P(H_{Col}) &= P_{const_1} + P(OT_1) + P(OT_2) \\ P(H_{Alr}) &= P_{const_2} + P(HV_{OD_{final}}) \end{aligned}$$

B.3 Constraint probabilities

Now we will introduce constraint probabilities. When looking at $P(H_{Alr})$ it is clear, that not all HVs at OD_{final} will trigger a false alarm. A false alarm is only triggered if the sensor is activated. This means either a) an OHV has activated it or b) both light barriers had misdetections before. So we introduce the constraint probability for $HV_{OD_{final}}$:

$$\begin{aligned} P_{constraint_1} = \;& P(OHV) + (1 - P(OHV))* \\ & *P(FD_{LB_{pre}}) * P(FD_{LB_{post}}) \end{aligned}$$

The first term in the sum refers to a) and the second to b). P(OHV) is the possibility for an OHV to be in the controlled area. So analogously to Sect. II-D we can now calculate the constrained probability for the cut set $HV_{OD_{final}}$:

$$P_{constrained}(HV_{OD_{final}}) = P_{constraint_1} * P(HV_{OD_{final}})$$

The cut sets $\{OT_1\}$and $\{OT_2\}$ are treated in the same manner. For all other cut sets we will include the constraint probabilities into P_{const_1} resp. P_{const_2}. In the end we get the following formulae[3] for the hazards probabilities:

$$\begin{aligned} P(H_{Col}) \;& = P_{const_1} \\ & +P(OHVcritical) * P(OT_1) \\ & +P(OHVcritical) * (1 - P(OT_1)) * P(OT_2) \\ P(H_{Alr}) \;& = P_{const_2} + (P(OHV) + (1 - P(OHV))* \\ & *P(FD_{LB_{pre}}) * P(FD_{LB_{post}})) * P(HV_{OD_{final}}) \end{aligned}$$

So these formulae respect information about how probable a “bad” behavior of the environment is.

C. Parameterized probabilities

We will now introduce parameterized probabilities. It is obvious, that the probability for an OHV to get stuck in a traffic jam such that it needs more than the maximum runtime of the timers P(OT_1) resp. P(OT_2), depends on the runtimes. In statistics there exist quite a lot of distributions[13], which describe such dependencies.

A good model for driving time of OHVs from LB$_{pre}$ to LB$_{post}$ and from LB$_{post}$ to the tunnel is normal distribution (mean time $\mu = 4$ minutes, standard deviation $\sigma =$ 2 minutes). $P_{OHV_{1/2}}(Time \leq T)$ denotes the probability for a driving time $\leq$ T. We can then calculate P(OT_1) in dependence of the runtime T1 of timer 1:

$$\begin{aligned} & P(OT_1)(T1) = 1 - P_{OHV_1}(Time \leq T1) \\ & where \\ & P_{OHV_1}(Time \leq T) := \frac{1}{\int_0^\infty exp(-\frac{(x-\mu)^2}{2\sigma^2})dx} \int_0^T exp(-\frac{(x-\mu)^2}{2\sigma^2})dx \end{aligned}$$

In the same manner parameterized probabilities for P(OT_2)(T2), P($HV_{OD_{final}}$)(T2), and P($FD_{LB_{post}}$)(T1) are calculated. We now substitute the parameterized formulae into the formulae for the probabilities of the hazards making them parameterized as well:

$$\begin{aligned} P \quad & (H_{Col})(T1, T2) = P_{const_1} + P(OHVcritical)* \\ & *(P(OT_1)(T1) + (1 - P(OT_1)(T1)) * P(OT_2)(T2)) \\ P \quad & (H_{Alr})(T1, T2) = P_{const_2} + *P(FD_{LB_{post}})(T1)* \\ & *(P(OHV) + (1 - P(OHV)) * P(FD_{LB_{pre}})) \end{aligned}$$

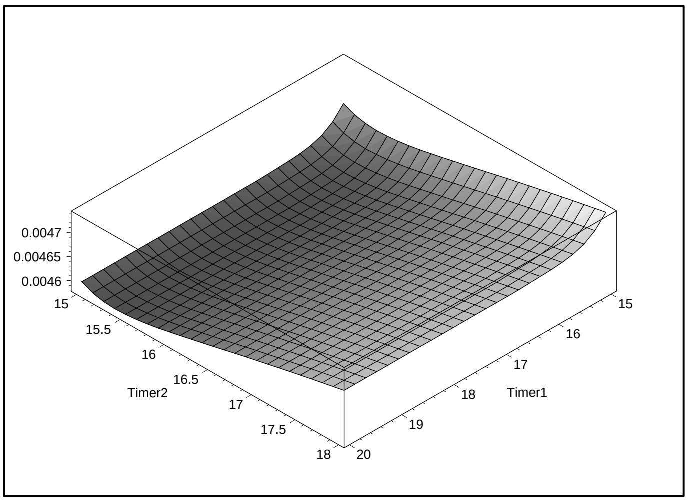

Fig. 5. The cost function around its minimum

Now we can easily calculate the probabilities for each hazard for an arbitrary choices of the timer runtimes. But to do optimization we need to formulate a connection between the two contrary hazards. Cost functions form this link.

C.1 Cost function

The cost function describes the relative importance of each hazard. In our example a good guess for the costs is, that collisions cost roughly 100.000 times the money a false alarm costs. So we get the cost function as the weighted sum of hazards probabilities:

$$\begin{gathered} f_{cost}(T1, T2) := \\ P(H_{Col})(T1, T2) * 100000 + P(H_{Alr})(T1, T1) * 1 \end{gathered}$$

The variables T1 and T2 represent the runtimes of the timers. $P(H_{Col})(T1, T2)$ resp. $P(H_{Alr})(T1, T2)$ denote the probabilities defined in the last paragraph.

C.2 Results

Figure 5 shows the cost function—plotted against the runtime of timer1 and timer2—around the minimum. Closer examination yields optimal parameters for the timer runtimes of approximately 19 resp. 15.6 minutes for timer 1 resp. 2. This is much less than the initial “guesses” of 30 minutes of the safety engineer and results in an improvement of about 10% in false alarm risk, while the risk for collision does not change (less then 0.1%). It also shows, that timer 1 may be chosen more conservatively than timer 2.

This result was somewhat surprising. Why does the false alarm risk only decrease by 10%, although the detection system is now only active half the time than before? And why is the dependency of the risk not symmetric in the free parameters? To answer this question we make use of the fact, that parameterized probabilities allow us to also examine the system in different working environments. We introduce an additional parameterized probability in the

[3] P(OHV critical) denotes the probability of an OHV driving towards west-tube or mid-tube

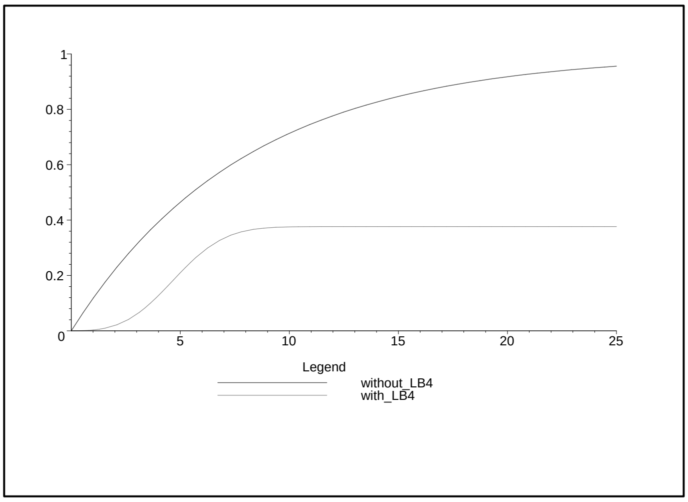

Fig. 6. Probability of false alarms, if an OHV is driving correctly

system - the rate of correct driving OHVs. This allows us to answer the question: How does the control scale if the traffic — especially the number of OHVs — increases. Figure 6 (graph "without LB4") shows the probability of a false alarm against the runtime of timer 2 assuming that an OHV is in the controlled area.

This result is alarming. It means, that even with the suggested, reduced runtime of 15.6 minutes for timer 2 more than 80% of the correct driving OHVs will trigger an alarm[4]. This makes the complex control system almost obsolete, as the same result could have been achieved much easier (for example by only using one light barrier and lock the tunnel whenever an OHV is detected), and is a major design flaw in the control.

The reason for this phenomenon can also be discovered by very precise informal analysis of the system (but neither we nor the engineers were aware of it before we did this analysis). When looking at the design of the system in figure 4 it becomes clear, that all countermeasures taken to avoid false alarms only aim at avoiding false alarms triggered by false detections of one of the light barriers. This is the main source of false alarms of the old height control of the former three tubes (there were no ODs at all as every OHV had to be stopped anyways).

But now the scenario is different. If OHV traffic increases, then the OD_{final} sensor will almost constantly be activated and this sensor will ultimately decided if an alarm has to be triggered or not. So the risk of a false alarm is then basically the probability of a FD at OD_{final} or a HV there. While we have some influence on the failure probability, our influence on high vehicles is limited. Although road regulations require trucks, buses and vans to drive on the right lane, some drivers always ignore this rule! The only possibility to reduce the the risk is to reduce the activation time of the sensor, but this is not feasible, as a runtime of less than 10 minutes will make the risk for a collision unacceptably high.

This obvious design flaw has neither been discovered by formal safety analysis with model checking nor by the engineers. The problem is, that the rigorous formal methods only give qualitative answers. The problems here lies in the logics. A statement can either be true or false, not possibly true[5]. Formal FTA showed that a false detection of OD_{final} is a critical single point of failure. Quantitative FTA did not discover the flaw, as normally no environmental constraints are considered there. It was only discovered, when we tested the concept of constrained probabilities on this case study and we were surprised by the results in the beginning as well. Because we introduced the frequency of correctly driving OHVs as a free parameter, we just had to take a look at the graph to discover the flaw.

The solution to this problem is fairly simple. It can be solved by introducing a new light barrier right in front of the entrance of tube 4 and use it to stop timer 2 (a counter for OHVs in zone 2 is then needed as well). The result of this light barrier is, that the mean activation time of OD_{Final} is now the mean time of an OHV needs to pass Zone2. The effect of this light barrier in the same scenario is shown in figure 6 (graph "with LB4"). The system will still ring the bell for a very high number($\approx 40\%$) of correct driving OHV, but this design can be implemented without major changes. A much better solution would be to install the light barrier at OD_{final}. This would lower the false alarm rate to approx. 4% of the OHVs, because the detector OD_{Final} would then only be critical, while an OHV passes the light barrier or the light barrier has a FD. However, it is not clear if this is physically possible with the current road layout and available light barriers.

V. Limitations and future work

It is our experience, that the results of this analysis depend a lot on how well the statistical model reflects reality. This problem can not be circumvented, but it's only a question of math and engineering to find good guesses for probabilities and distributions.

However, even if the statistics are not very elaborate, safety optimization can help by giving a rough estimation about how important the different parameters are. Furthermore constrained probabilities allow to "try" different statistical assumptions and at least get an estimation for the worst case.

An interesting point for future work is to combine safety optimization and formal fault tree analysis [12][15]. This could make it much easier to find the difficult constrained probabilities as a formal proof of a cause-consequence relationship also implicitly contains information about the necessary environmental inputs. In fact the formalization of INHIBIT-gates yield necessary condition for the constraints. So it is a promising idea to *collect* all INHIBIT-gates along the paths from the fault tree root to the leaves of a cut set. The result should be a formal description

[4] For a runtime of 30 minutes it is more than 95%.

[5] At least standard formal methods. Logics like probabilistic CTL (PCTL), combined with probabilistic model checking[1] can also address such problems.

of the constraints necessary to make the primary failures force the hazard's occurrence.

However, the problem of dependent probabilities stays. This is especially true for constraint probabilities. They are often dependent. Formal methods might help there, too. For example if logical implication of two constraints ($A \rightarrow B$) can be shown and the constraints are described by predicate logic formulae, then P(A) is an upper bound for P(B). It would be interesting to investigate, if the dependency can be expressed for other cases as well. The natural next step is to drop the restriction to predicate logic constraints and examine constraints with duration or temporal orderings.

Another open end for research is of course to find limitations which other distributions and cost functions are necessary in real-world applications. This also raises the question in which cases the resulting optimization problem stays solvable. An interesting connection is to reduce the whole optimization problem to a problem of stochastic programming, which is a branch of mathematical optimization that deals with probability distributions.

We found, that in practice the results of different analysis methods like formal verification, FTA, quantitative FTA and safety optimization may be used as input data for each other, so an integrated methodology will make safety analysis a lot more systematic, easier and cheaper. This is—together with intuitive tool support—also a key feature for the possible application in industrial practice.

VI. Conclusion

Our experience is that it is important to combine different techniques for safety analysis. This is because different methods not only examine different aspects of the system, but also give contrary views [5].

Safety optimization is one such technique. The idea of safety optimization is as simple as promising: do a fault tree analysis of the systems hazards, use statistical distribution for failure probabilities, estimate the costs of each hazard with a cost function and do mathematical optimization. The result will be an optimal configuration of the system with respect to the examined hazards.

While traditional safety analysis does not assess the problem of usability and trustworthiness, such issue may be considered with parameterized probabilities as well. It tunrs out, that analyzing the system in different working environments and analyzing the effect of free parameters can be treated in the same way. But here the focus lies in the analysis, as it is only of theoretical interest what the optimal working environment of the system might be.

Safety optimization is an extension of fault tree analysis. It extends the quantitative aspects of FTA. Together with formal FTA [12][11] which extends the qualitative aspects and allows to prove that the cause-consequence relationship between primary failures and hazards is correct, this analysis is of very high significance.

We illustrated the benefits of the new method with a real world case-study: the Elbtunnel in Hamburg. The safety analysis of the height control for the Elbtunnel has shown the benefit of the combination of all these methods. Formal verification has shown a design flaw, which resulted in a safety gap. Fault tree analysis identified critical failure and yielded minimal cut sets for quantitative analysis. Quantitative safety analysis, showed the importance of different failure modes and gave upper bounds for hazard probabilities. Safety optimization yielded optimal configuration values for free parameters and discovered a major design flaw. This flaw has neither been detected by the system engineers nor by a formal verification of correctness. Safety optimization has made the system safer, led to design improvements and increased overall system quality.

In conclusion, we find that examining only hazards and estimating their probabilities is not enough. It is rather important to examine all hazards of a system in parallel, its intention and the planned working environment to get a good impression of it. A combined approach of traditional safety analysis, formal methods and mathematics can accomplish this. Such an integrated approach is being developed within the ForMoSA research project [3][11].

References

[1] Christel Baier, Edmund M. Clarke, Vassili Hartonas-Garmhausen, Marta Z. Kwiatkowska, and Mark Ryan. Symbolic model checking for probabilistic processes. In *Automata, Languages and Programming*, pages 430–440, 1997.

[2] J. Fragole J. Minarik II J. Railsback Dr. W. Vesley, Dr. Joanne Dugan. *Fault Tree Handbook with Aerospace Applications.* NASA Office of Safety and Mission Assurance, NASA Headquarters, Washington DC 20546, August 2002.

[3] FORMOSA – formal models and safety analysis, 2001. http://www.informatik.uni-augsburg.de/swt/formosa/.

[4] A. H. G. Rinnooy Kan G. L. Nemhauser, editor. *Optimization*, volume Vol 1. Elsevier Science Publishers B.V, 1989.

[5] E.G. van den Blieck J.L. Rouvroye. Comparing safety analysis techniques. *Reliability Engineering & System Safety*, 2002.

[6] N. Leveson. *Safeware: System Safety and Computers.* Addison Wesley, 1995.

[7] Nancy G. Leveson. A new approach to system safety engineering. Aeronautics and Astronautics Massachsetss Institue of Technology.

[8] David G. Luenberger. *Linear and nonlinear programming.* Addison-Wesley Publishing Company, 1989.

[9] K. L. McMillan. *Symbolic Model Checking.* Kluwer Academic Publishers, 1990.

[10] F. Ortmeier, W. Reif, G. Schellhorn, A. Thums, B. Hering, and H. Trappschuh. Safety analysis of the height control system for the Elbtunnel. In *SafeComp 2002*, pages 296 – 308, Catania, Italy, 2002. Springer LNCS 2434.

[11] F. Ortmeier and A. Thums. Formale Methoden und Sicherheitsanalyse. Technical Report 15, Universitt Augsburg, 2002. (in German).

[12] G. Schellhorn, A. Thums, and W. Reif. Formal fault tree semantics. In *Proceedings of The Sixth World Conference on Integrated Design & Process Technology*, Pasadena, CA, 2002.

[13] Klaus Schürger. *Wahrscheinlichkeitstheorie.* R. Oldenbourg Verlag, 1998.

[14] N. Storey. *Safety-Critical Computer Systems.* Addison-Wesley, 1996.

[15] A. Thums and G. Schellhorn. Formal safety analysis in transportation control. In R. Slovak and E. Schnieder, editors, *Proceedings of the Workshop on Software Specification of Safety Relevant Transportation Control Tasks*, Braunschweig, Germany, 2002. VDI-Verlag.

[16] W. E. Vesely, F. F. Goldberg, N. H. Roberts, and D. F. Haasl. *Fault Tree Handbook.* Washington, D.C., 1981. NUREG-0492.