SAND2021-3548C

Risk-Informed Access Delay Timeline Development

Douglas M. Osborn, Dusty Brooks, and Andrew Thompson*
Sandia National Laboratories
PO Box 5800, MS-0789, Albuquerque, NM 87185-0789
dosborn@sandia.gov; dbrooks@sandia.gov; andthom@sandia.gov

ABSTRACT

The U.S. Department of Energy Office of Nuclear Energy's Light Water Reactor Sustainability Program is developing a new method to modernize how access delay timelines are developed and utilized in physical security system evaluations. This new method utilizes Bayesian methods to combine subject matter expert judgement and small performance test datasets in a consistent and defensible way. It will also allow a more holistic view of delay performance that provides distributions of task times and task success probabilities to account for tasks that, if failed, would result in failure of the attack. This paper describes the methodology and its application to an example problem, demonstrating that it can be applied to access delay timeline analyses to summarize delay performance using subjective and objective information.

Key Words: Physical Security Risk

1 INTRODUCTION

Using the current methods, access delay timelines rely on reported data from tests where possible, and on SME judgement to help fill in any blanks that exist in the testing. This data is generally reported using a single time rather than distributions. When a distribution is desirable and data is available, a conservative approach tends to be to assume that the minimum time reported from testing is the mean, and then a minimum and maximum are assumed to be +/- 50% of this mean in a symmetrical triangular distribution. This assumes that the fastest test time was close to the average and that there is no chance that actual task time could be less than 50% of the mean or more than 150% of the mean. However, it is common for performance testing teams to be highly experienced with breaching tools and methods since it is rare to bring in untrained individuals for testing. Additionally, the tests frequently include a set of testing where the attackers have no knowledge of the barrier, and then a second one where engineers will walk the attack team through details and point out the parts of the design that are likely the weakest. As a result, the fastest times recorded are often more than likely much shorter than the true mean and more representative of a time in the tail of the distribution. Compounding these issues, it isn't uncommon for analysts to use the shortest times from these distributions to be conservative. These triangular distribution assumptions lead to highly conservative reported timelines that do not provide a good indication of the true risk.

In many cases, these paths require a significant number of tasks, each coming with its own risk of delays or failures. Tool failure, fatigue, injury, or even simply the fact that it is unlikely that every step will be executed nearly perfectly are not accounted for. In other areas of physical security, potential of failure is addressed, such as a sensor failing to detect an adversary or a response force being rendered unable to complete their mission, but this has not been extended to the access delay analysis historically.

Typically when using this data in a timeline, it is reported as a single point, rather than as a distribution that accounts for all of the test data, as well as neglecting being able to account for SME judgement on the

* Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2020-XXXX

shortest possible test completion time if every sub-task went as smoothly as possible. This project was the result of looking into ways of integrating SME judgement with a handful of data points from testing and utilizing that to provide a wider picture of the probability distributions associated with task times as well as full timelines.

1.1 Benefits to Shifting to a Risk-Informed Timeline Method

While the historical methods outlined above have helped to increase security at a range of facilities over the past decades, utilizing new methods could provide a broader understanding of the security posture at a facility. The key driver for considering a change in methods is to provide a more accurate assessment of the true delay times, and in order to achieve this task times need to be viewed as distributions rather than relying on a single data point to better gauge the difficulty of a task. In addition, the opportunity to consider the probability of successfully completing a task is also a key element that is currently not considered. If a timeline is very unlikely, it may be prudent to focus protective measures on a different timeline that an adversary is much more likely to complete, even if it happens to be a longer timeline.

In analysis of physical security systems, we see these probabilities used in a wide range of system effectiveness calculations. In detection, the probability of detection along with a confidence interval is estimated using testing and applied for each sensor in the system. For assessment and response, human factors data is used to estimate the probability that a threat will be properly assessed and to predict response force reactions. These tests are not destructive and can be repeated many times, allowing for data sets large enough for frequentist methods. To date, this type of criteria is not used in delay, and the single-point style timeline is used as ground truth to compare everything against. This does help to provide a conservative approach, but significant data is often neglected. This can cause security professionals to focus on specific scenarios and lose sight of other areas that may also benefit from additional attention.

The goal of this effort is to investigate how delay timelines can be presented as distributions as well as to develop methods with strong statistical foundations to generate the distributions that are needed for this style of analysis. One of the challenges to implementing a statistical approach is the unavailability of relevant and high-quality data. Due to the high cost and destructive nature of many delay tests, we are required to utilize small delay data sets to generate distributions, which limits the types of statistical methods that will be useful. Bayesian analysis is a natural discipline for this, as it is can provide more useful results with less data by combining subjective information (subject matter expert judgement) with objective measurements (existing data).

This effort focuses on the implementation of Bayesian methods for combining subjective and objective information in a way that is meaningful for timeline analysis. However, there are additional aspects to this type of analysis that are not addressed. As an example, this methodology does not explicitly account for interaction effects (correlations) in the probabilities of task completion throughout the timeline. These effects, as well as concerns such as fatigue or environmental effects, are included in the subject matter expert (SME) judgement component of the methodology but not in the stage in which distributions are updated with data. These limitations do not outweigh the benefits, but should be considered as methods mature and are applied to more complex problems

1.2 Applicability of Bayesian Methods

Bayesian statistics is a statistical theory centered around the concept of Bayes' rule, which follows from the basic rules of conditional probability. This section describes the basic concepts and applicability to timeline analysis. The reader is referred to [1] and [2] for much more comprehensive treatments of the material. For a parameter θ and data set D, Bayes' rule states that the probability distribution on θ given the data, $P(\theta|D)$, can be calculated using:

$$P(\theta \mid D) = \frac{P(\theta \mid D)\, P(\theta)}{P(D)} \tag{1}$$

It can be helpful to conceptualize this equation by considering a single point. Consider an example where the parameter θ is the probability of an adversary succeeding at a task. We want to use both data and our expert knowledge to draw inferences about the distribution of the probability of adversary success, θ. In other words, we want to use the data and expert judgements to calculate the probability that $\theta{=}\theta^*$ for all possible values θ^*. For any given value θ^*, we write Eq. 1 as:

$$P(\theta = \theta^* \mid D) = \frac{P(D \mid \theta = \theta^*)P(\theta = \theta^*)}{P(D)} \tag{2}$$

Notice that we are interested in θ, but the denominator $P(D)$ is independent of the value of θ. Because of this, the equation is often simplified to a statement of proportionality:

$$P(\theta = \theta^* \mid D) \propto P(D \mid \theta = \theta^*)P(\theta = \theta^*) \tag{3}$$

The term $P(D)$, called either the marginal likelihood or evidence, is essentially a normalizing constant that ensures the final distribution is a proper probability distribution, so it is unimportant to consider for understanding the concept.

We can think of each component of Eq. 3 as the answer to a specific question, as in Table I. The posterior distribution is what we want to understand, the likelihood is calculated from the data, and the prior characterizes our subjective state of knowledge independent of the data. The power of Eq. 3 is that it says we can calculate the distribution on the probability of adversary success if we (1) can characterize our prior state of knowledge about adversary success as a distribution, and (2) we collect data on adversary success.

Table I. Components of Bayes' theorem and the associated terminology

Expression	Concept	Term
$P(\theta{=}\theta^*\|D)$	How likely is it that the probability of adversary success is θ^* given the data D that we have observed?	Posterior
$P(D\|\theta{=}\theta^*)$	If we knew that the probability of adversary success was θ^*, how likely is it that we would have observed the data that we have?	Likelihood
$P(\theta{=}\theta^*)$	How likely is it that the probability of adversary success is equal to θ^*?	Prior

The prior and the likelihood together form our understanding. If, for example, we think the probability that $\theta{=}\theta^*$ is high based on our SME knowledge, the value of the prior at θ^*, $P(\theta{=}\theta^*)$, will be high. But if the data is inconsistent with this information, the likelihood $P(D|\theta{=}\theta^*)$ will be low. Because the prior and the likelihood are multiplied together, both must be high for the posterior to assign high probability to θ^*.

There is some additional nuance in this relationship between data and belief within the Bayesian framework – the size of the data set matters immensely. The effect of the beliefs characterized by the prior distribution diminishes as the size of the data set grows. This means that the methodology leverages SME judgements to augment small data sets but relies upon it less when it is not needed.

2 IMPLEMENTATON OF BAYESIAN METHODS FOR TIMELINE ANALYSIS

This work considers a seven-step method for applying Bayesian methods to timeline analysis:

1. Identify a group of SMEs who will participate. There may be different SMEs for different tasks and SMEs should only provide values for tasks within their area of expertise.
2. Elicit minimum, best estimate, and maximum values from the experts on the success probability for the task. If SMEs are uncomfortable assigning minimum or maximum values, another common approach would be to ask for 1% and 99% bounds.

Mathematically, the final distribution on adversary success probability for the task will include the full interval (0,1), so whether the experts provide minimum/maximum values or low/high percentiles has a negligible effect. Either phrasing for the elicitation has the effect of defining a subinterval of (0,1) with high density. The significance in phrasing the question is in maintaining consistency with the way SMEs are comfortable thinking about the problem.

3. Choose a triangular distribution based on these elicited values.

   One option is to use the average of the elicited minimums to define the lower bound, the average of the elicited best estimates to define the mode, and the average of the elicited maximums to define the upper bound. However, this method may not always be appropriate and there are multiple options. The phrasing "a triangular distribution based on" is intentionally technically vague because the way this is accomplished is a subjective decision point for the analyst. See Section 3 for a discussion on alternative methods.

   Note that the triangular distribution step may be unnecessary. The SME judgements could be used directly to estimate a beta distribution in the next step, skipping this step. However, because triangular distributions are the most common distributions currently applied for uncertainty in timeline analysis, this methodology was constructed to apply either to new analyses, or to previous analyses which used triangular distributions. There is an added benefit to including this step in that triangular distributions are more intuitive than beta distributions, which gives the analyst multiple options (see Section 3) for heuristically selecting a prior with minimal statistical expertise.

4. Choose a beta distribution, $beta(a_{prior},b_{prior})$, to approximate the triangular distribution.

   This can be done by performing maximum likelihood estimation (MLE) though other fitting methods may be used as well. In the classical definition of a beta distribution, a_{prior} and b_{prior} have domains with no upper bound. However, the larger the values of these parameters, the more weight the prior has in determining the posterior – this situation is particular to this usage of the beta distribution and is not true more generally. So, the methodology used should allow for constraints on the parameters. See Section 4 for further technical discussion of this step.

   Replacing the triangular distribution with a beta extends the domain of the triangular distribution to (0,1) so experts do not need to have absolute confidence in their lower/upper bounds; values arbitrarily close to 0 or 1 will be included with low density even if the expert judgements do not extend near 0 and 1. It also smooths the distribution. This beta distribution, $beta(a_{prior},b_{prior})$, is the prior for that task.

5. Collect data or perform testing on the task. There are two types of data that can be used to update the prior.

   i. Trial data – this data takes the form of successes (or failures) out of trials, k successful attempts at a task out of n attempts. This may include trials of the same task, or tasks that are of similar difficulty according to SME judgement.

   ii. Probability data – this data takes the form of a probability estimate (or multiple probability estimates). As with the trial data, this may include probability estimates for the specific task under consideration or estimates for similar tasks.

6. Update the prior with the collected data. The method used to perform this update will depend on the type of data that was collected.

   i. When trial data is used, the posterior has an analytical solution obtained via conjugacy with the binomial distribution. For k successes in n trials, the posterior is $beta(a_{prior}+k,b_{prior}+n-k)$ [1]

   ii. When probability estimate data is used, the prior should be updated using Bayesian software. This situation may occur when the data is obtained through a literature search, for example. One source

may provide details of a modeling and simulation activity that resulted in an estimate, p_1, for probability of success. Another source may summarize testing in which they only provide p_2 (calculated as k/n) without reporting k or n. These sources would both be useful data points to include in a timeline analysis, even though they provide less information than would be ideal.

7. Propagate the posterior distribution through the event tree that describes the timeline using Monte Carlo analysis methods. Depending on timeline complexity, this can be implemented in most software that can perform random sampling, multiplication, and if/then logic or more specific software with explicit event tree and fault tree implementation.

This effort focuses on applying Bayesian methods to the probability of adversary success because inclusion of success probability is in itself an advancement over standard methods. Applying uncertainty analysis on these probabilities with Bayesian methods Additionally, most relevant data for this type of analysis will be in the form of trial data, which allows an analytical solution for the posterior without Bayesian software or expertise. The method can be expanded to derive uncertainty distributions on task completion time; however, this will likely require implementation in Bayesian software because there may not be an analytical solution for the posterior. Triangular distributions were used for task completion times in the example analysis and were defined as above, except that experts were asked to provide minimum, best estimate, and maximum task completion times rather than success probabilities. This demonstrates the implementation of task time uncertainty into the overall methodology but leaves Bayesian development of this component to future work.

3 COMBINING ELICIATIONS

Section 2 mentions one method for combining expert-elicited values: average the values over the SMEs. However, there are cases where this method can lead a final distribution that does not characterize the combined knowledge well or that sacrifices too much conservatism. This section discusses, heuristically, different philosophies on combining these judgements and how the decision can affect the final result. It is important to remember that this part of Bayesian timeline analysis is the subjective component – there is no one correct method to combine the SME judgements, but the analyst can choose the method purposefully to reinforce analysis priorities. This discussion focusses on combining elicitations into a triangular distribution, which is used in this methodology as a starting point for forming a beta prior. However, recall from Section 2 that an approach which directly forms a beta prior from the SME judgements without this intermediate triangular distribution step is equally valid.

One option for forming the triangular distribution, instead of using the average of the SME bounds to define the domain as in Section 2, is to use the minimum of the SME lower bounds and the maximum of the SME upper bounds. This would extend the triangular distribution's domain, which may be considered conservative with respect to the amount of uncertainty included (i.e., this characterizes a belief that there is high uncertainty). However, this method could be nonconservative with respect to risk if it shifts significant mass to lower probabilities of adversary success. Similarly, the mean of SME best estimate judgements on a linear scale may be a reasonable measure of central tendency (as is the median). On a log scale, the median is a reasonable measure of central tendency whereas the mean is typically highly conservative. If, for example, the distribution of SME judgements spans many orders of magnitude, the values in the highest part of this range will dominate the arithmetic mean and can result in a 'best estimate' that the majority of experts judge to be unreasonable.

The method used to form the triangular distribution can be chosen after studying the SME judgements. The balancing of expert opinions towards the center could be a desired result; one of the motivations for using a set of multiple experts is to prevent an analysis from being driven by one extreme opinion. However, the domain of this distribution is small compared to the expanse of values covered by all expert opinions together. It may then be preferable to use one of the methods that calculates the lower and upper bounds as the minimum and maximum, respectively, to better characterize the high level of uncertainty. These

distributions would reflect a belief that there is high uncertainty, but the most likely value is still roughly in the center of expert judgements.

4 FORMING A BETA PRIOR

As mentioned in Section 2, it can be important to constrain the magnitude of the beta distribution's parameters in the prior. This is because the posterior for $beta(a_{prior},b_{prior})$ with data in the form of k successes out of n trials is $beta(a_{prior}+k,b_{prior}+n-k)$. If $a_{prior} \gg k$, then $a_{prior}+k \cong a_{prior}$; meaning the experimental data would have almost no effect on the first parameter of the posterior. Similarly, if $b_{prior} \gg n-k$, then $b_{prior}+n-k \cong b_{prior}$ and the experimental data would have almost no effect on the second parameter of the posterior.

The choice of upper bounds for a_{prior} and b_{prior} is subjective and should be chosen to reflect how strong the prior belief is. The classical "uninformed" beta priors (Bayes, Jeffreys, and Haldane) all have $a_{prior},b_{prior} \leq 1$. This constraint works well for uniformed priors but does not always work well for informed priors (such as our prior informed by SME elicitations) because it severely limits the shape of the beta distribution.

One method for limiting the sizes of a_{prior} and b_{prior} is to increment the bounds until an acceptable fit to the SME judgements (whether using the judgements themselves or fitting to the triangular distribution) is reached. In keeping with the concept that the first parameter represents successes and the second parameter represents trials minus successes, it makes sense to also impose the restriction that $a_{prior} \leq b_{prior}$; if the belief represented by the distribution is that the success probability is low. If the belief is that the probability of success is high (most trials result in success), then $b_{prior} \leq a_{prior}$ could be used as a restriction.

Alternatively, the analyst can base upper bounds by thinking of the SME judgements as trials. We may decide based on the experience level of our SMEs that each judgement is worth one trial, or 'one and half trials', etc. If we have four SME judgements to use in defining our prior and consider each to have the same weight as half of a trial, our SME judgements collectively represent two trials. This would result in the constraint $a_{prior},b_{prior} \leq 2$. We could also restrict the values of a_{prior} and b_{prior} relative to each other as discussed in the previous paragraph. This type of heuristic reasoning is not uncommon for beta prior distributions.

Bounding a_{prior} and b_{prior} (as described in the last paragraph) by numbers greater than 1.0 increases flexibility in the beta distribution shape compared to the uninformed beta distributions. However, it is nonetheless a limitation that may exclude the 'best fit.' This beta distribution is meant to be similar to, but should not be expected to closely fit, the triangular distribution obtained using SME judgements. To judge whether the beta prior is acceptable, it should be compared to the SME judgements.

5 FULL PHYISCAL SECURITY SYSTEM ANALYSIS INTEGRATION

In addition to being able to aid in standalone analysis of delay timelines, developing these distributions can also be integrated into tools that analyze the full physical protection system. Rather than a task having a single time assigned, each run would be able to sample a time from a distribution. This provides a more realistic model of an adversary actions rather than assuming the conservative assumption that all tasks are accomplished in an optimal time. The low end of this will still provide the same conservative estimate, but the accumulation will better demonstrate the real-world time that one would expect to see the timeline completed in.

A more complete understanding can have many benefits on understanding the physical protection system, as well as helping to identify key tasks where additional delay would have the most impact on

buying down overall risk. Currently, it is not uncommon for much of the focus to be placed on the shortest timelines, even if they would require an extremely unlikely sequence of events to occur. In these situations, there may be timelines that are longer but more consistent that would be better places to focus to buy down overall risk. By utilizing this method on multiple timelines and adjusting the distributions at key points, the effects of system upgrades can be compared, and optimal places to utilize funds can be identified.

As with any delay analysis, ensuring that appropriately trained individuals are involved will be key to the success of this implementation. This should include delay SMEs that would normally be involved with timeline analysis but should also include a statistician or analyst with experience in Bayesian methods to ensure a sound statistical approach is used. Though the overall process can be described prescriptively in high-level steps as done in this paper, the execution in practice should be done carefully to ensure consistency between SME judgements and their probabilistic representations.

Caution is also warranted with respect to the relationship between timeline structure and the associated task probabilities. The number of sequential tasks in a timeline determines the number of probabilities multiplied together to get the final timeline success probability. This means that the result is highly sensitive to decisions that, on a task-by-task basis, may seem insignificant.

For example, consider a timeline with sixty consecutive tasks, as was the case for the timeline used in this study. Assume that two SMEs separately assign a success probability to each task and both SMEs think success is likely for each task. One SME characterizes "likely" as a 0.99 probability of success for each task; the other SME characterizes "likely" as a 0.95 probability of success for each task. Qualitatively, this seems reasonable and there is not much of an appreciable difference for an individual task; both judgements state that the adversary will most likely succeed the task.

However, when this slight difference in judgements is propagated through the sixty-task timeline, the difference is magnified at each step. According to the first SME, the probability of success for the full timeline is $0.99^{60} \approx 0.55$. According to the second SME, the probability of success for the full timeline is $0.95^{60} \approx 0.05$. That is an order of magnitude difference in the result, even though both SME's judgements seem reasonable and similar at first glance.

The proposed methodology requires distributions instead of the point estimates like those in this example because uncertainty is known to be high. However, this also means that the uncertainty in the final timeline success probability will be even more magnified. This is less of a concern with less complex timelines. If the timeline in this example had just five tasks instead of sixty, the two SME estimates for final timeline success probability would be 0.95 and 0.77 respectively. Thus, the way that the timeline is structured and the granularity with which tasks are defined has significant implications for the mathematics of the problem. Dividing a timeline into more tasks will tend to drive the estimate of timeline success probability down, so care should be taken during timeline construction to avoid dividing tasks into subtasks unnecessarily.

The methodology proposed in this paper should be implemented cautiously and thoughtfully with these considerations in mind. This may impact timeline construction, by limiting the complexity of the timeline itself. Or, this may impact where in the timeline the methodology is applied. It may be preferable to use conservative estimates for some tasks, while applying the method to quantify uncertainty only on selected tasks where it makes sense to do so; a combination of conservative assumptions and uncertainty estimates provides a more complete picture than conservative assumptions alone.

6 CONCLUSIONS

Shifting from traditional methods of delay timeline analysis, in which a single value is used to represent a task's duration, to one that provides probability distributions for both task duration and likelihood of success will help analysts and facility operators to better understand the risk associated with various potential attack pathways. This will help facilities prioritize funding to the areas that will buy

down the most risk rather than being driven mostly by the pathways with the shortest credible times. This is expected to help facilities improve their security posture without driving up cost excessively and may open opportunities to expand design-basis threat considerations.

The methodology presented in this paper is meant to enable security risk managers to begin this transition into a more holistic type of delay timeline analysis. The statistical tools are simple, and the methods are structured to utilize analytical solutions for ease of implementation. However, the method will be most useful when the statistical tools are married to timeline construction so the benefit can be gained through targeted, thoughtful use of the method; wholesale application to any timeline should be avoided.

Extensions to this methodology can be developed that will enable analysts to explicitly account for correlations between tasks, as well as factors like progressively decreasing adversary performance through fatigue, or progressively increasing performance through practice. Flexibility can be increased by transitioning from Bayesian statistical models with analytical solutions, such as how the beta distribution was applied in this paper, to a broader population of models with solutions that can be obtained with sampling. The characterization of uncertainty within the methodology could also be advanced through SME training on statistical distributions and elicitation facilitation to enable experts to specify distributions with more flexibility.

7 ACKNOWLEDGMENTS

We would like to thank the subject matter experts from Sandia's Access Delay and Structural Assessment department for their time and effort in helping to develop timelines for this effort and those who provided valuable feedback while developing the methods outlined in this paper. We would also like to thank Sandia's C.J. Hartwigsen, W. Gary Rivera, and Greg Wyss for both their technical review and exceptionally insightful discussions.

8 REFERENCES

1. J. K. Kruschke, *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*, Elsevier Inc., (2015).
2. A. Gelman, J.B. Carlin, et al., *Bayesian Data Analysis*, CRC Press, (2013).