Global Information Assurance Certification Paper

Modeling Security Investments With Monte Carlo Simulations

GIAC GWAPT Gold Certification

Author: Dan Lyon, danlyon@mac.com
Advisor: Stephen Northcutt

Accepted: September 16, 2014

Technical leaders not only guide technical activities but are also required to manage stakeholders to drive investment in key architectural components. Models and simulations can be leveraged to represent and communicate the impact of architectural decisions. This paper describes how to use published information on the cost of security data breaches to build a model using Monte Carlo simulations and Crystal Ball with Microsoft Excel. An application of the model is illustrated by picking three example controls on application security from the Critical Controls for Effective Cyber Defense.

1. Introduction

Technical leaders and architects are frequently the interface from sponsors and management into projects. According to Microsoft, architects represent the executive sponsor and need to understand the business case for investments (Nema, 2006). But the translation also must work the other way, as the architect must be able to convince the executive team of the need for architectural investments that may not have an obvious business case. This is evidenced through The Open Group's TOGAF framework for developing enterprise architecture (Hornford, 2011). One of the guidelines provided through TOGAF is that architects must perform Stakeholder Management, which involves capitalizing on positive messaging to stakeholders. This requires speaking in the language of the business stakeholders.

When businesses evaluate investment options, they use cost-benefit analysis to help identify return on investment (PMBOK Guide, 2013). An architect should provide the financial analysis as one of the criteria for project selection. This can be difficult with architecture and in the case of security investments can be even harder because it is about minimizing risks (How to Build the Business Case for Enterprise Architecture, 2011).

An approach to bringing the cost-benefit analysis to architecture creation is to build a financial model of the expenses or cost savings involved. The value proposition is one aspect of the architecture that can be weighed when making decisions on technical options, as reported by the Software Engineering Institute (Nord, et al, 2003).

One of the problems with security is determining how much benefit one gains from various security investments. This is the realm of risk management and contains many unknowns. Bruce Schneier notes that traditional models like Annualized Loss Expectancy fall apart when it comes to modeling rare and expensive events (2008). Schneier also notes that the models for security return on investment are good in theory but not valuable in practice. One of the reasons cited is because of differing opinions on how much things may cost which cannot be argued. One method to moving away from opinions and arguments is to use a model with simulations. Monte Carlo simulations allow for repeatedly using random values in the calculation of equations, which allows for a range of uncertainty (Maas and McNair, 2009). Using a range of values with

Dan Lyon, danlyon@mac.com

probabilities is much easier to reach agreement on that single values that may be used in Annualized Loss Expectancy.

Since Schenier's post on this, much more data has become available. Ponemon Institute's 2014 Cost of A Breach report published data that can be used for building a model to simulate breach event impact.

In this paper, an example model will be developed that allows an architect to communicate with stakeholders in the traditional business language of money. This paper is not intended to dispute Schneier's reality of modeling events and forecasting return on investment, however it does show a tool available to security architects that may be useful in managing stakeholders.

According to the Department of Defense Systems Management College (2001), the four major benefits of modeling and simulation are cost avoidance, increased quality, expedited schedule and cost savings. This paper shows a method to forecast cost avoidance using published data on cost reduction of information security breaches. Systems engineers and architects can use such forecasting to make decisions on which security controls provide the most benefit to the business.

2. Building a Model

When a financial model is created, it typically consists assumptions, inputs, calculations based on inputs and assumptions, and outputs (Fairhurst, 2009). Each of the four aspects must be defined to create a model that can evaluate financial benefits of various security controls. A sample model was built using Oracle Crystal Ball and Microsoft Excel. Maas and McNair (2009) highlighted the use of Monte Carlo simulations as a way to include uncertainties when evaluating a product's business case.

The set of assumptions used for the examples here will be taken from the 2014 Ponemon Institutes Cost of Data Breach report. First, a model was built from the 2014 values. Second, calculations were run on the generic model to illustrate the use of the model. In the third step sets of sample controls are used with development costs to showcase how the model can be used to help make decisions. Fourth and last, the model is updated for multiple years of data and compared against a second source for validation.

Dan Lyon, danlyon@mac.com

The methodology and the limitations of the Ponemon Institute are not described nor validated by this effort. The information is used to create a model. If a model like this is implemented, it is the responsibility of the creator to validate all assumptions and ensure that the data used is suitable for the purpose.

Table 1 – Breach Data Models

Assumption On Breach	2014 Value
US Average cost per record	201
Max Cost per record	359
Min Cost per record	100
Most Likely Cost per record	
Breach Cost per record	0
US Average number of breached records	29087
Max Records breached	100000
Min Records breached	2415
Most Likely Records Breached	29087
Min Reduction in Breach Cost	1
Max Reduction in Breach Cost	42
Most Likely Reduction in Breach Cost	21

2.1. Model Overview

The first items to be defined are the assumptions. In the sample model, assumptions are created using the 2014 Cost of Data Breach Study. The model depicted in

Table 1 – Breach Data Models is described, including assumptions and distributions used. It should be highlighted that all of the assumptions and inputs described in the next sections are variables that need to be modified for any given simulation. This section will address the generic makeup of the model, and the following section will show the application of the model to different security solutions. Lastly, the model is updated to reflect multiple years of data and incorporate a second source.

2.2. Model Assumptions

The assumptions in the model are based upon breach cost per record. The Ponemon Institute data was examined for each of the rows listed. The first assumption

captured in table is the cost per breached record to the organization. The model has a minimum cost per record, a maximum cost per record, and a most likely cost per record. The values extracted from the Ponemon report are $100 minimum, $359 maximum, and $201 average cost per record.

The second assumption used is the number of records breached. This assumption also needed minimum, maximum and most likely number of records. In this value set, the most likely was assumed to be 10% of the maximum number.

The first and second assumptions are evaluated by simple multiplication to create a breach cost forecast. The breach cost forecast represents the total cost to the organization from the breach. Following the cost forecast, any reduction per record is accounted for by the next set of assumptions.

Minimum, maximum and most likely values were taken from the Ponemon studies. Note that the maximum reduction value was calculated by adding four separate data points. These four data points were identified in the Ponemon report as: strong security posture, incident response planning, business continuity management, and CISO appointment. For simplicity and comparison, the assumption was made that the four data points broken out in 2014 represented the maximum reduction. The 2013 Ponemon report listed a single value for cost reduction per record, and is the rationale why a single data point was used instead of four separate data points. The minimum reduction in breach cost was determined to be $1 because that is the lowest value that has an impact on the cost of the breach. The most likely reduction was assumed to be 50% of the maximum value.

2.3. Model Calculations

The calculations used are those provided by Oracle's Crystal Ball plugin for Excel. Crystal Ball works by using Monte Carlo simulations to produce data based on assumptions. Crystal Ball assumptions are defined by using any of the provided probability distributions. For this model, the triangle distribution was repeatedly chosen because it allowed for a range of values (min, max) that can be held constant and a most likely value that can be varied based on different inputs.

An example of the triangle distribution is shown in Figure 1. The 2014 Breach Cost Per Record value is described by the minimum, maximum and most likely values shown by Figure 1 – Breach Cost Per Record. When the simulations are run, for each simulation the value of Breach Cost per Record will be selected randomly from this distribution graph. When Crystal Ball is run, the number of simulation runs is configurable, and you can select large numbers (5000 or more) to get realistic values.

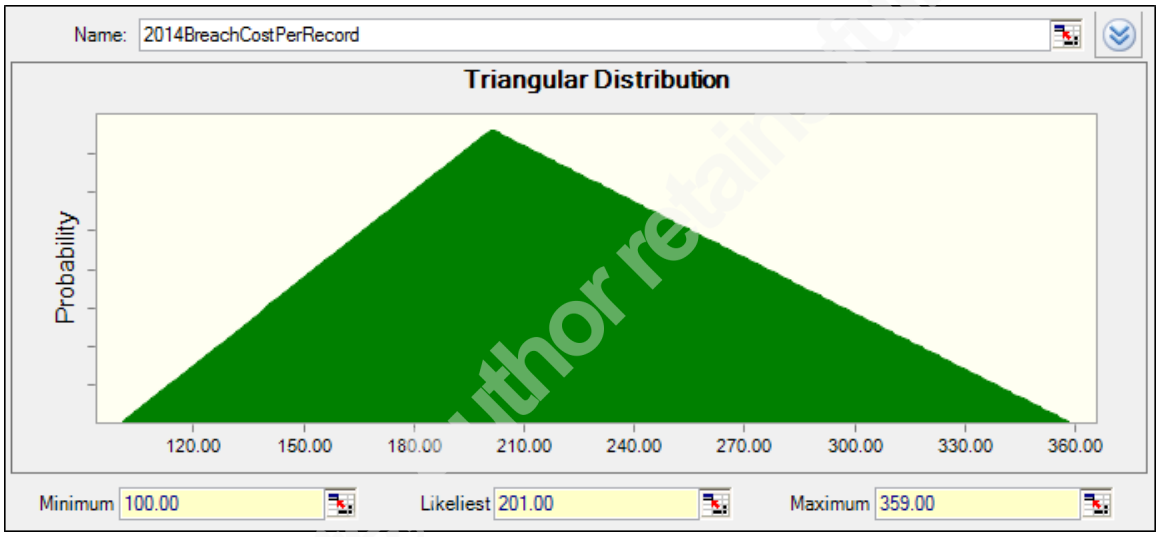


Figure 1 – Breach Cost Per Record

The next value that must be provided is the number of records breached. The values that were chosen from the 2014 Ponemon Institute report are used in a triangle distribution for the minimum and maximum. The most likely number of records to be breached in this example is chosen as 10% of the maximum. This most likely value would need some analysis by technical staff to determine an appropriate value. It could be selected based on knowledge of the system such as a typical number of records transferred. The graph is provided in Figure 2 - 2014 Number of Records Breached.

Dan Lyon, danlyon@mac.com

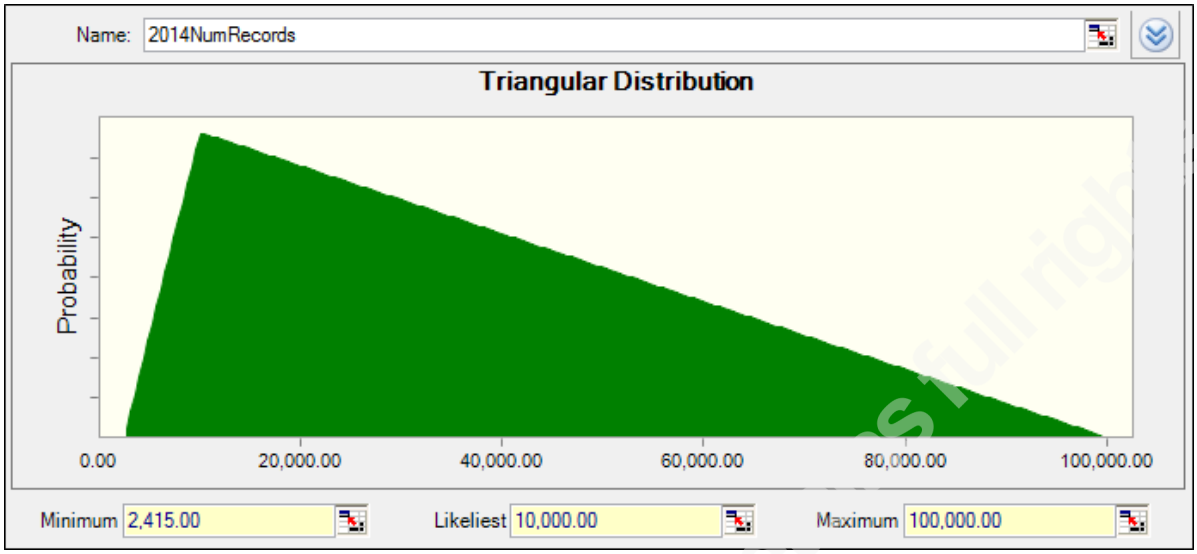


Figure 2 - 2014 Number of Records Breached

The benefit of security is implemented as a cost reduction per record breached because that was the data available. The value chosen for the maximum was from the 2014 Ponemon report, while the minimum was selected as one dollar because that was determined to be the lowest amount of benefit a security controls may provide. The most likely was chosen to be the midpoint. All values can be seen in Figure 3 - 2014 Cost Reduced per Record.

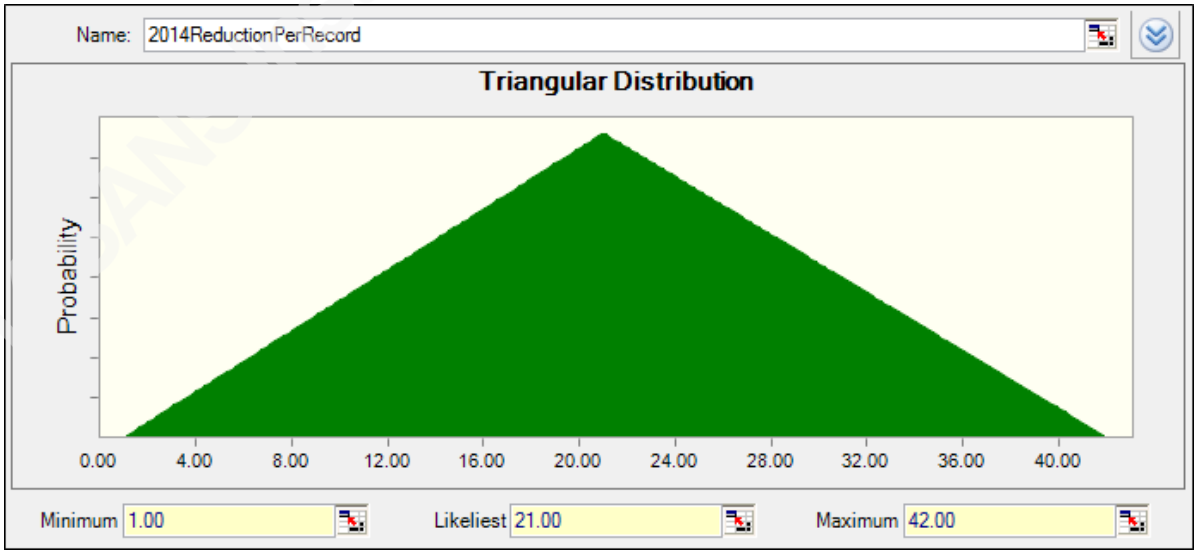


Figure 3 - 2014 Cost Reduced per Record

Dan Lyon, danlyon@mac.com

2.4. Model Outputs – Breach Cost

The model output that is desired is a forecast on what a breach may cost, and a probability distribution associated with the values. This is accomplished in Crystal Ball through definition of a Forecast. The mathematical equation used to forecast is simply the

Number of Records Breached * Cost Per Record Breached

In Crystal Ball, the formula is placed in the Excel cell, and then the forecast is calculated from the formula. When the formula includes Crystal Ball assumptions, the formula is calculated using the assumptions. The forecasted graph is displayed in Figure 4 - Breach Cost Forecast which shows that a breach may cost anywhere from $0 to over $20 million. This is what the model forecasted, but further analysis is required. In the 2014 Ponemon report, the maximum breach cost reported was $5.85 million, and therefore values above that are questionable and should be addressed.

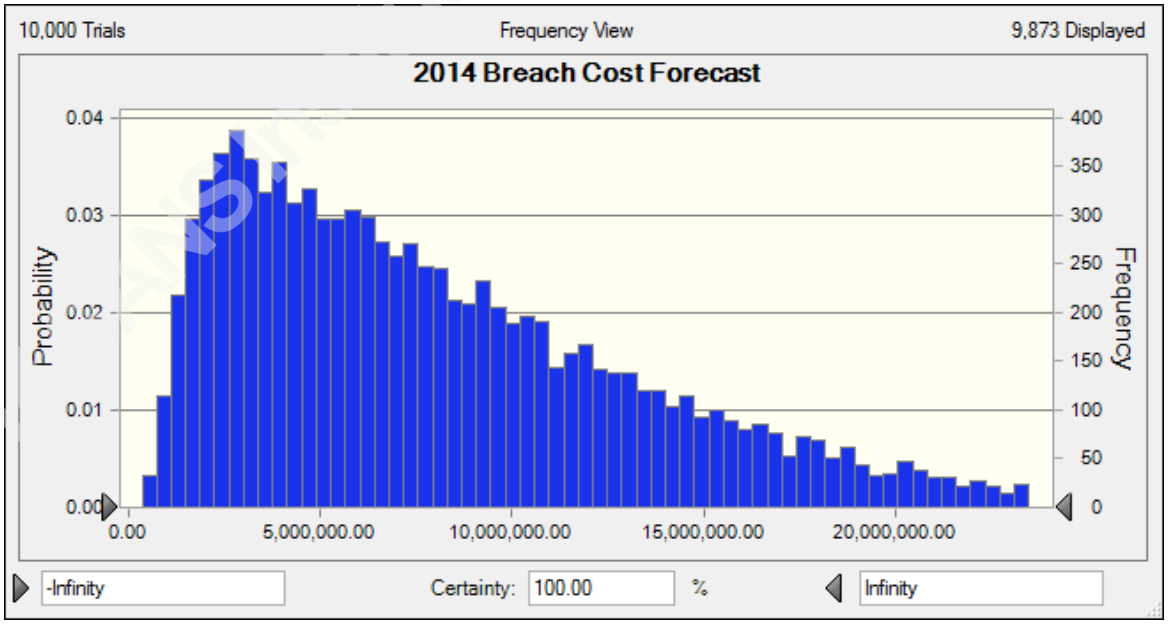


Figure 4 - Breach Cost Forecast

Dan Lyon, danlyon@mac.com

The distribution plot provided by Figure 5 – Limited Breach Cost Forecast shows the projected cumulative probability plot for Breach Cost Forecast after accounting for the maximum breach cost noted in the Ponemon report. To reflect a maximum value, the model was set to filter out all values above six million. This can be used to set a maximum on the cost impact of a breach.

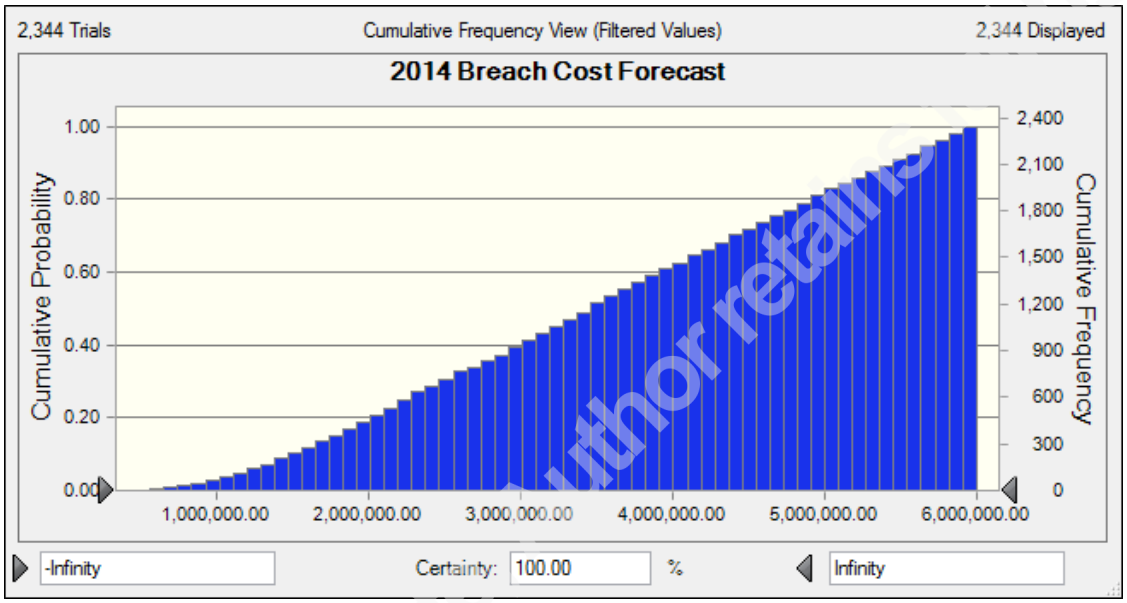


Figure 5 – Limited Breach Cost Forecast

2.5. Model Outputs – Cost Reduction

The forecasted breach cost value is useful only when it drives action in terms of identifying security controls that could be implemented. The next step required is accounting for the reduction in breach cost that security controls provide in the event of a breach. The Breach Cost After Reduction equation is:

Number of Records Breached * (Cost Per Record Breached - Breach Cost Reduction Per Record)

The resulting forecast for breach cost after reduction is displayed in Figure 6 - Reduced Breach Cost with a cumulative probability graph, again with results filtered to remove any values over six million.

Dan Lyon, danlyon@mac.com

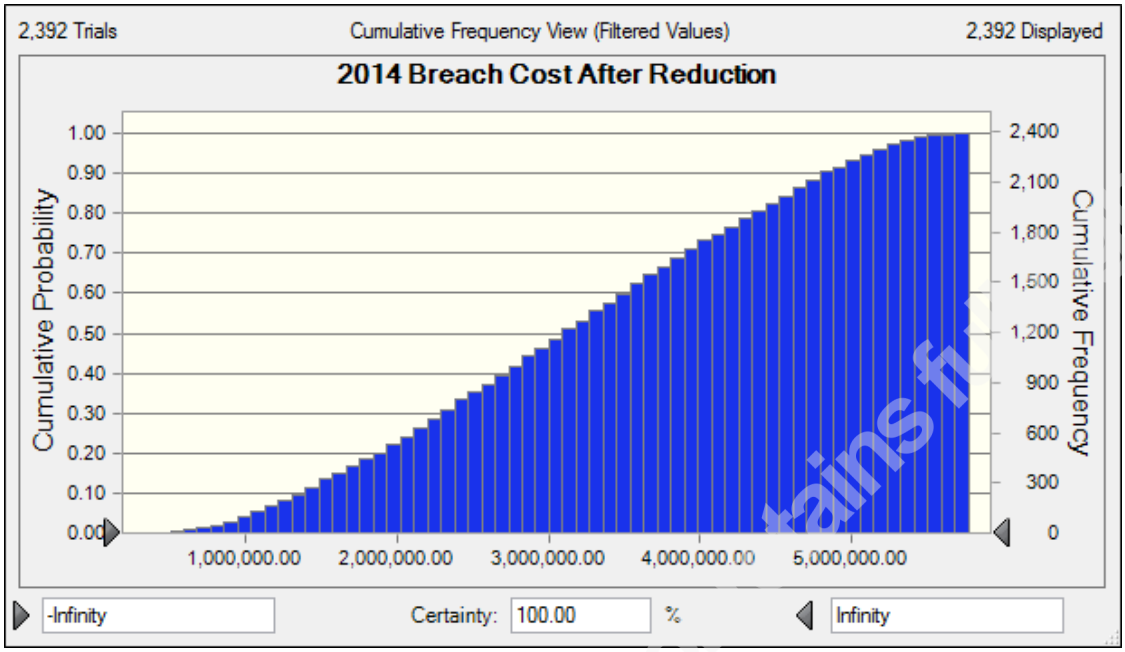


Figure 6 - Reduced Breach Cost

The cost savings from an increased security posture can be calculated by calculating the difference between the cost before and the cost after reduction is applied. The potential range of cost savings is displayed in Figure 7 - Cost Savings. A model with predicted cost savings can be used to compare architectural choices by varying the model inputs and assumptions. An application of the model is presented next to illustrate its use.

Dan Lyon, danlyon@mac.com

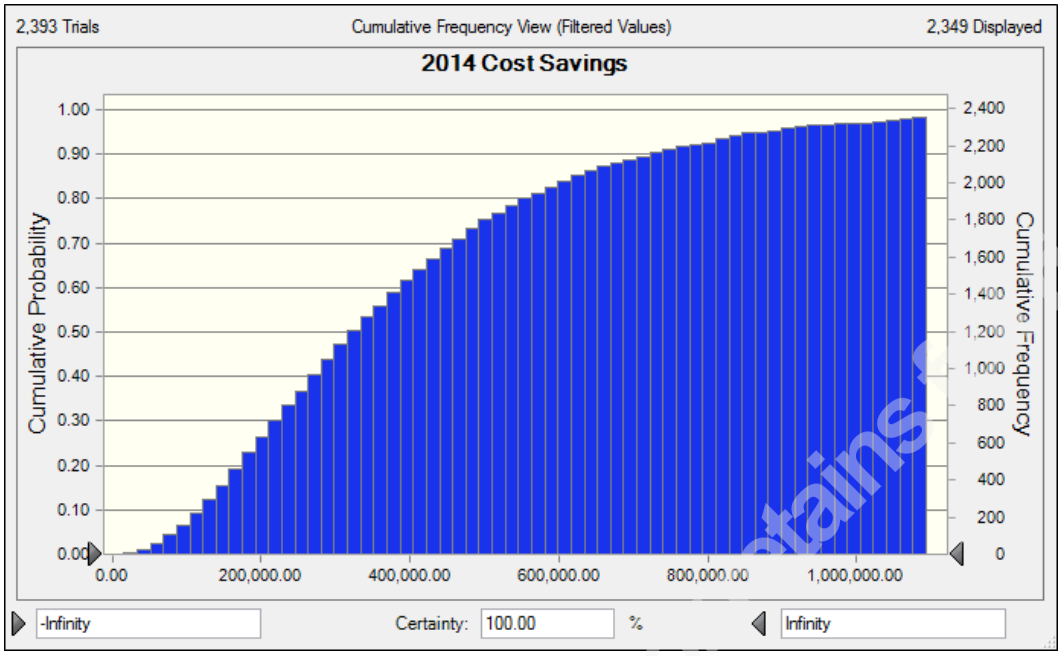


Figure 7 - Cost Savings

3. Application of Model

For an example of how to use the model, three sample application software security investments were chosen from the Critical Security Controls for Effective Cyber Defense (2014). The three examples analyzed are: implementing a web application firewall, architectural upgrades to improve security, and training for the software development organization.

First, a common rating scale is established for experts to rank the value of each of the potential upgrades. A five point scale of low, medium-low, medium, medium-high and high is used to estimate the security posture of the system after each of the upgrades. Each of the estimated scale items can be translated to a cost reduction per record, and the values are shown in **Table 2 - Security Posture Conversion**.

Table 2 - Security Posture Conversion

Estimated System Security Posture	Equivalent Value ($)
Low	1
Medium-Low	11

Dan Lyon, danlyon@mac.com

Medium	21
Medium-High	31.5
High	42

Assume a web application firewall would cost $15000 to purchase and implement (Barracuda Web Application Firewall, 2013). Staff estimated it would improve the security posture of the system from low to medium-low. A medium-low security posture is translated to a cost reduction of $11 per record. For the model, the range of cost reduction that can be achieved from this investment can be estimated to range from $1-11. A most likely value of $5 can be chosen of in the middle of the range to give a triangle distribution.

In the second example, assume that architectural upgrades have been identified that could make significant improvements into the system security. Assume that experts have estimated that the upgrades would cost about $200,000 and that security posture improvement would move to high. Architecture upgrades can be modeled with cost reduction values of $30 (minimum), $42 (maximum) and $37 (most likely).

For the third upgrade, training for the software development organization has been identified. But training by itself will not increase the organizations posture, so some additional effort must be expended to improve the existing implementations. Assume that the cost of training and follow on implementation effort is $200,000. The improvements are thought to increase the security posture to medium. Training benefits can be modeled with cost reduction values of $15 (minimum), $26 (maximum) and $20 (most likely).

Each of the estimated projects has uncertainty around the costs because projects frequently run over budget (Bloch, Blumberg & Laartz, 2012). Cost overrun or under run can be modeled in the same manner as described above. In this example, a triangle distribution using min-max-most likely values is used. The most likely value is the estimate received for each implementation, and the min and max values are a percentage of the most likely. The application firewall is projected to be within 5% of actual cost, because it is a smaller effort. The larger projects of architectural upgrades and training are estimated to be within 20% of final cost.

Dan Lyon, danlyon@mac.com

Analyzing the benefits of the various options is reduced to calculating the cost reductions with the chosen values, and subtracting the implementation costs. Using Crystal Ball, this is done by establishing forecasts. Table 3 - Per Record Cost Reduction shows the Crystal Ball assumptions for each of the options. Assumptions were established with a triangle distribution using the values in the table.

Table 3 - Per Record Cost Reduction

Value, in $ per record breach	**Web App Firewall**	**Architecture Upgrades**	**Training**
Min Reduction	1	30	15
Max Reduction	11	42	26
Most Likely Reduction	5	37	20

The final model is displayed in Figure 8 - Crystal Ball Mode. The cost savings for each of the options are calculated to show the differences in benefit between the three options. In this example, the training and the architecture upgrades cost the same amount, however when the cost savings are compared, the architectural upgrades are a better investment because it offers significantly more potential for cost savings in the event a breach occurs. The architecture investments have the potential for saving more than either of the options. The three options resulting cost savings are compared in an overlay graph shown in

Dan Lyon, danlyon@mac.com

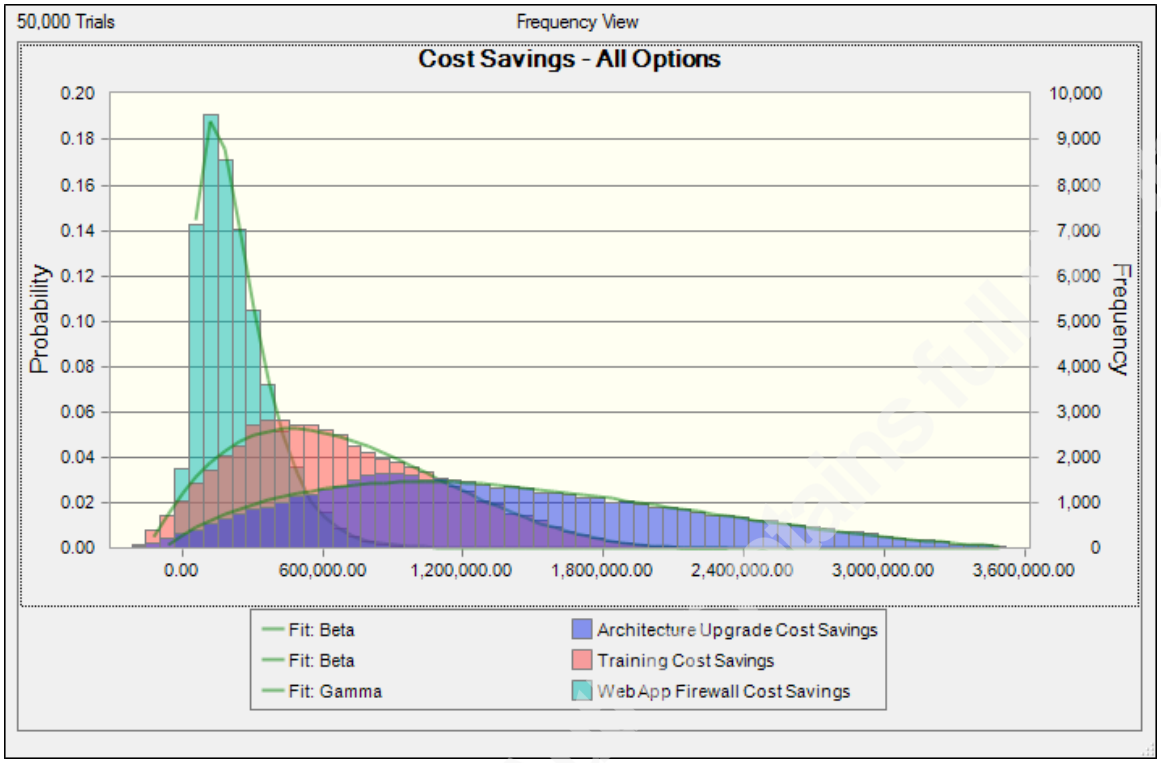
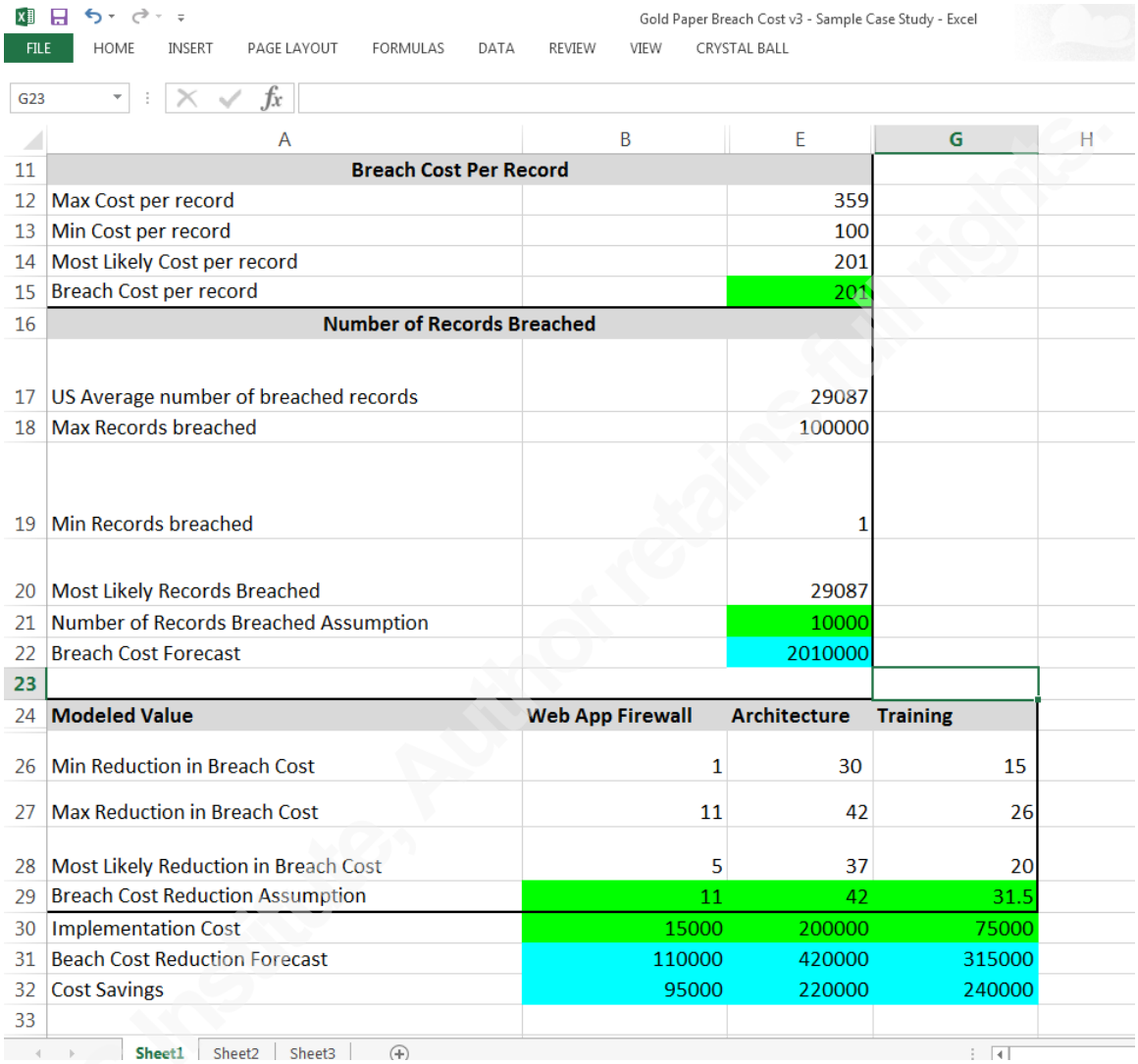


Figure 9 – Cost Savings Options Comparison.

Dan Lyon, danlyon@mac.com



Gold Paper Breach Cost v3 - Sample Case Study - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW CRYSTAL BALL

G23

	A	B	E	G	H
11	**Breach Cost Per Record**				
12	Max Cost per record		359		
13	Min Cost per record		100		
14	Most Likely Cost per record		201		
15	Breach Cost per record		201		
16	**Number of Records Breached**				
17	US Average number of breached records		29087		
18	Max Records breached		100000		
19	Min Records breached		1		
20	Most Likely Records Breached		29087		
21	Number of Records Breached Assumption		10000		
22	Breach Cost Forecast		2010000		
23					
24	**Modeled Value**	**Web App Firewall**	**Architecture**	**Training**	
26	Min Reduction in Breach Cost	1	30	15	
27	Max Reduction in Breach Cost	11	42	26	
28	Most Likely Reduction in Breach Cost	5	37	20	
29	Breach Cost Reduction Assumption	11	42	31.5	
30	Implementation Cost	15000	200000	75000	
31	Beach Cost Reduction Forecast	110000	420000	315000	
32	Cost Savings	95000	220000	240000	
33					

Sheet1 Sheet2 Sheet3

Figure 8 - Crystal Ball Model

Dan Lyon, danlyon@mac.com

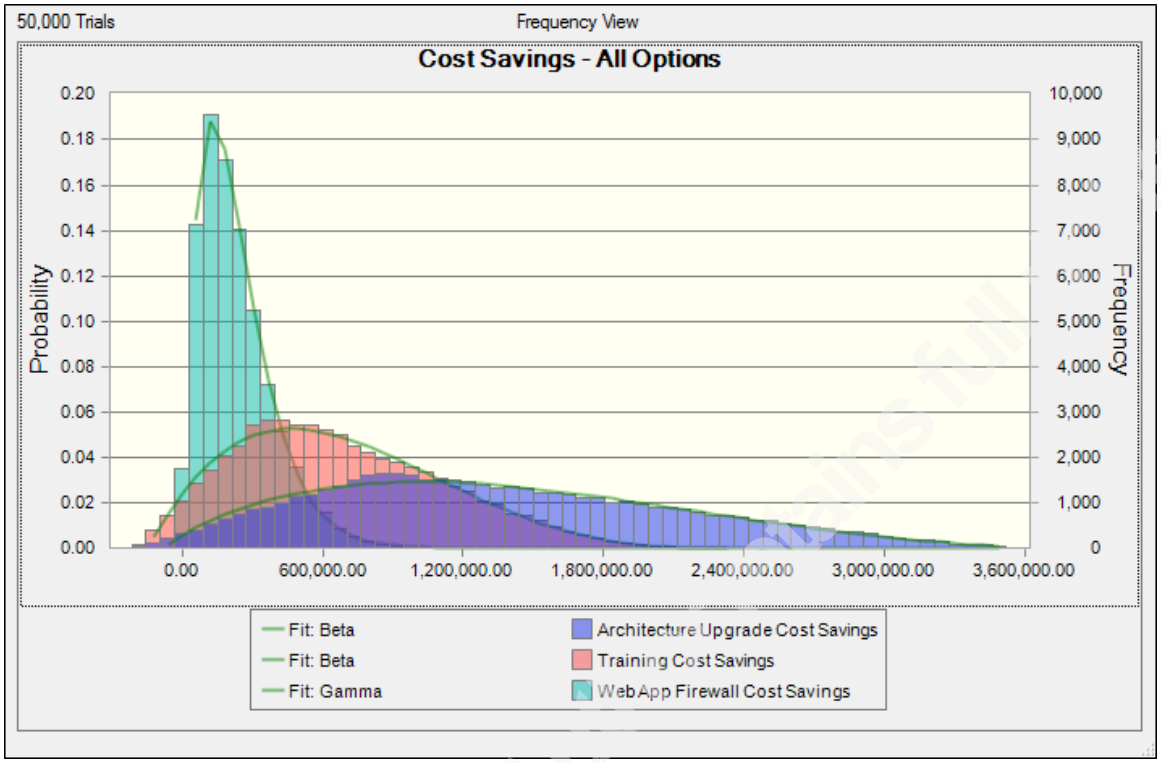


Figure 9 – Cost Savings Options Comparison

While the model created and discussed is simplified for discussion, the end result is that the benefit to security investments can be evaluated. When a cost-benefit analysis is required for various investment options, simulations can be used to provide the analysis and communicate the benefits of various investments.

One of the benefits to a model is the ease with which assumptions can be modified and tested. As mentioned previously, all distributions chosen in the above example are triangle distributions; however changing the assumption from a triangle to a beta PERT distribution has a significant impact. The Beta PERT distribution is used exclusively for modeling expert estimates using a min, max and most likely value (ModelRisk Help). In this model, while the data provided is real, it could also be viewed as expert opinion. The main difference in between the Beta PERT distribution and the triangle distribution is that Beta PERT is four times more sensitive to the most likely value.

When the assumptions are changed to all use Beta PERT, the change in forecast is dramatic. The impact on the Cost Savings is shown in Figure 10 – Beta PERT Assumption Cost Savings. Under this assumption, a different decision may be reached about the value of each of the controls because the benefits are very different in terms of

cost savings impact. The difference highlights why it is critical to have validated assumptions for any model that may be created.

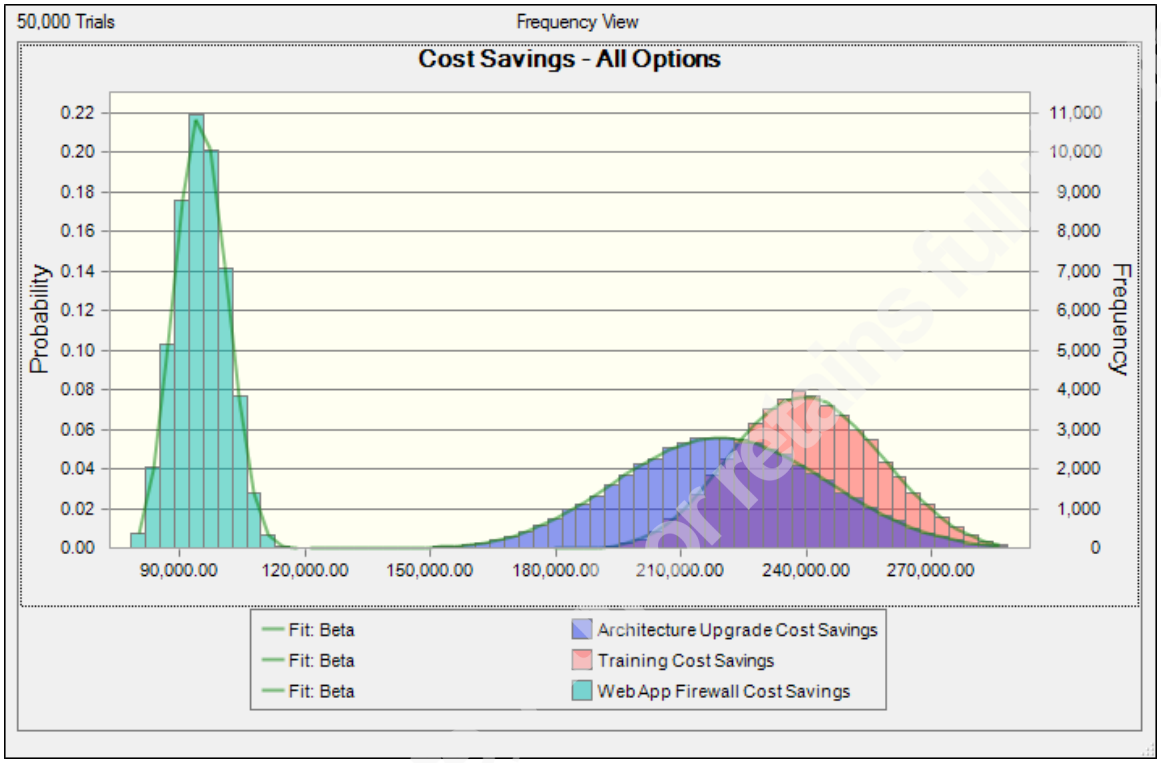


Figure 10 - Beta PERT Assumption Cost Savings

4. Adding Data to the Model

The model was described was built upon a single source of data. The model can be improved by incorporating more data. Multiple years of data are added to the average breach record cost, and a second source of information is evaluated to provide validation of model assumptions.

Crystal Ball has a time series predictor feature that allows forecasting of numbers based on historical data. After all years of data from the 2011 Ponemon report are used in a time series prediction as shown in Table 4 the forecast cost per record is about $190. Figure 11 shows the cost per record prediction from Crystal Ball. The model can now be updated with data based on 9 years instead of a single data point.

Table 4 - Average Yearly Cost per Record of Breaches

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013

Dan Lyon, danlyon@mac.com

Average cost per Record (U.S $)	138	182	197	202	204	214	194	188	201

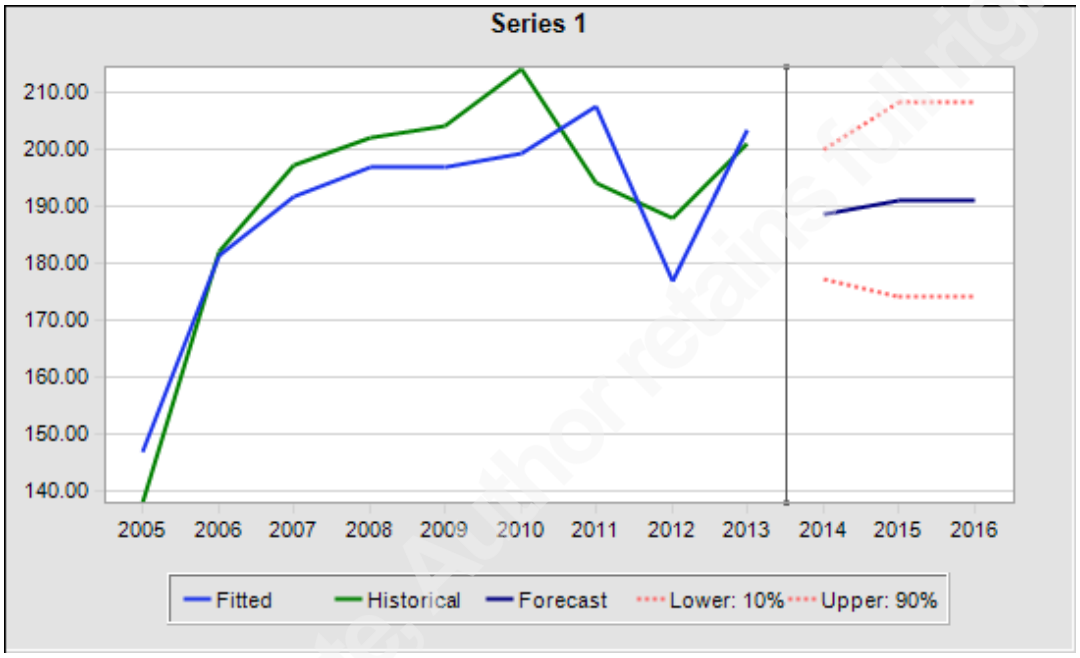
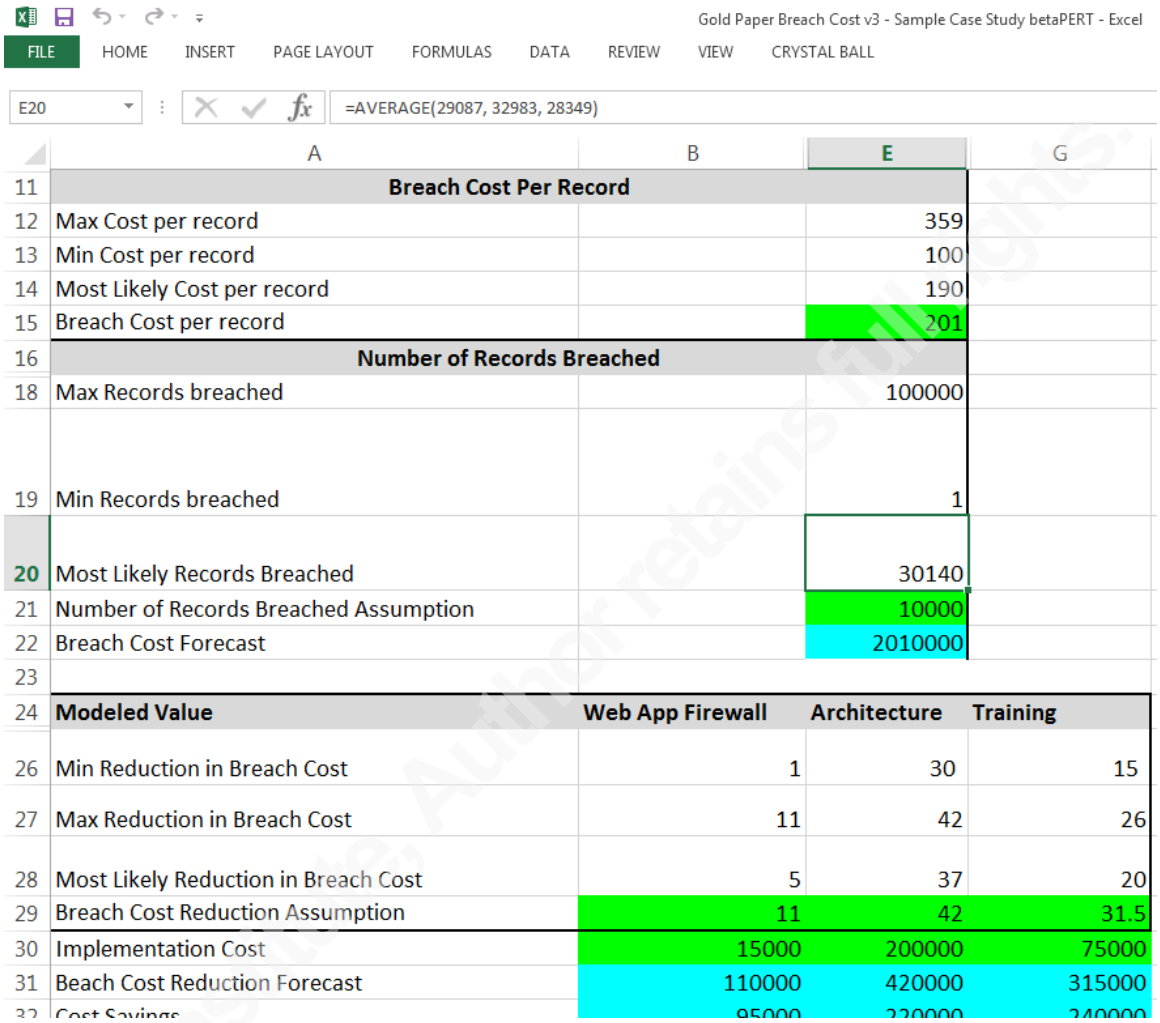


Figure 11 - Cost per Record Time Series Prediction

A second source of data breach cost evaluated was Navigant's Information Security and Data Breach Report from March 2014. Navigant's report shows an average cost per record of $188 for both 2012 and 2013. The time series value of $190 is supported by Navigant's analysis showing $188. Thus, the model is updated to use 190 as the most likely cost per record of a data breach.

Navigant also shows an average number of records breached as 32983 for 2013, compared to Ponemon's 29087 for 2013 and 28349 for 2011. Using these values, the most likely number of records breached was updated to the average value, 30140.

The model was updated using the averaged values, as shown in Figure…



Gold Paper Breach Cost v3 - Sample Case Study betaPERT - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW CRYSTAL BALL

E20 =AVERAGE(29087, 32983, 28349)

	A	B	E	G
11	**Breach Cost Per Record**			
12	Max Cost per record		359	
13	Min Cost per record		100	
14	Most Likely Cost per record		190	
15	Breach Cost per record		201	
16	**Number of Records Breached**			
18	Max Records breached		100000	
19	Min Records breached		1	
20	Most Likely Records Breached		30140	
21	Number of Records Breached Assumption		10000	
22	Breach Cost Forecast		2010000	
23				
24	**Modeled Value**	**Web App Firewall**	**Architecture**	**Training**
26	Min Reduction in Breach Cost	1	30	15
27	Max Reduction in Breach Cost	11	42	26
28	Most Likely Reduction in Breach Cost	5	37	20
29	Breach Cost Reduction Assumption	11	42	31.5
30	Implementation Cost	15000	200000	75000
31	Beach Cost Reduction Forecast	110000	420000	315000
32	Cost Savings	95000	220000	240000

Figure 12 - Model Updated for Multiple Years

The model forecasting cost savings for the three architectural options was rerun, and the updated cost savings is shown in Figure 13 – Multiple Years Predicted Savings. The values are not significantly different than those shown in Figure 10, showing that small changes to the model do not make large differences in the output.

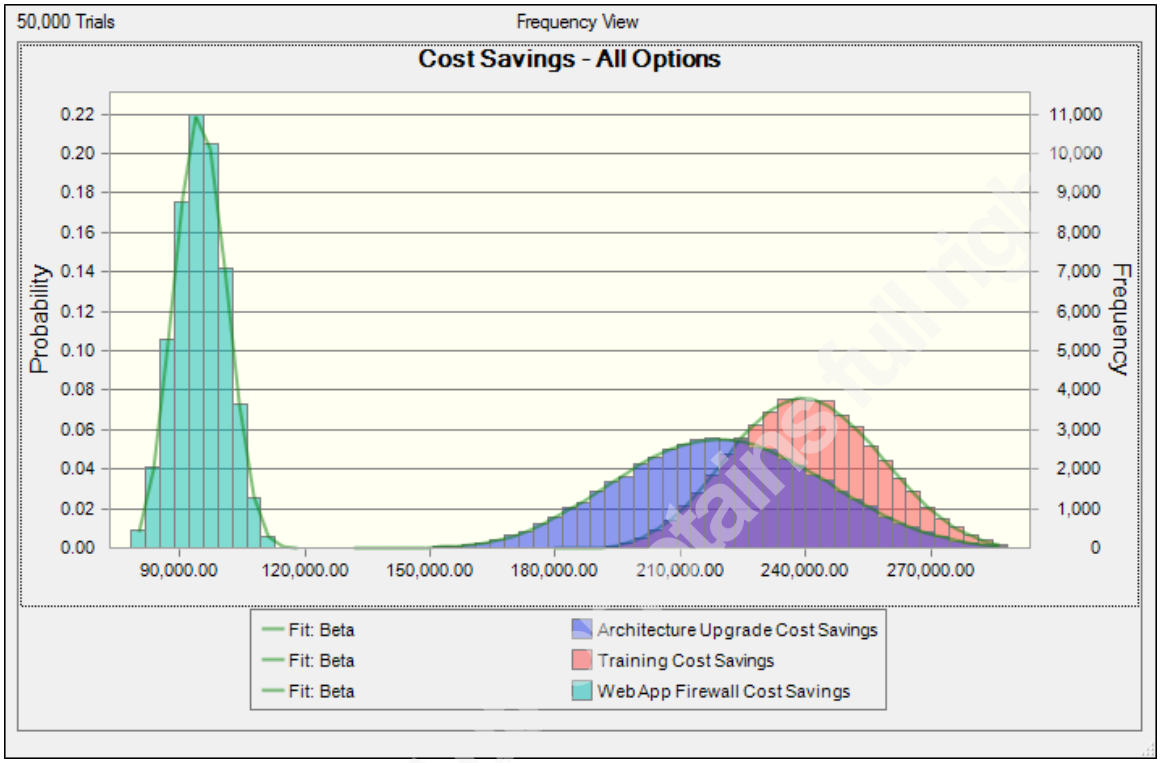


Figure 13 - Multiple Years Predicted Savings

5. Conclusion

Businesses must make financial-based decisions, and as Susan Landau wrote in IEEE Security & Privacy (2014) design choices are made with little understanding of the eventual impact. While evaluating the security benefits of alternative architectures is difficult, the model presented here is one method that architects can leverage to understand technology choices in a financial manner. Furthermore, the model can be used to communicate technical benefits with project sponsors. Another application of a model is to evaluate vendor claims using assumptions that have been created and reviewed by experts within the organization.

Dan Lyon, danlyon@mac.com

6. References

A guide to the Project Management Body of Knowledge (PMBOK guide), fifth edition (5th ed.). (2013). Newtown Square, Pa.: Project Management Institute.

Barracuda Web Application Firewall. (2013, November 1). Retrieved August 17, 2014, from http://www.scmagazine.com/barracuda-web-application-firewall-model-660/review/4039/

Bloch, M., Blumberg, S., & Laartz, J. (2012, October 1). Delivering large-scale IT projects on time, on budget, and on value. Retrieved August 17, 2014, from http://www.mckinsey.com/insights/business_technology/delivering_large-scale_it_projects_on_time_on_budget_and_on_value

Critical Security Controls for Effective Cyber Defense, version 5.0. (2014). Retrieved August 17, 2014, from https://www.sans.org/media/critical-security-controls/CSC-5.pdf

Department of Defense Systems Management College. (2001, January). *Systems Engineering Fundamentals*. Defense Acquisition University Press, Fort Belvoir, Virginia. Retrieved August 9, 2014 from http://ocw.mit.edu/courses/aeronautics-and-astronautics/16-885j-aircraft-systems-engineering-fall-2005/readings/sefguide_01_01.pdf

Fairhust, D. S. (2009, November 16). Six reasons your spreadsheet is NOT a financial model. *Blog*. Retrieved August 9, 2014, from http://www.fimodo.com/2009/11/six-reasons-your-spreadsheet-is-not-a-financial-model/.

Hornford, D. (2011). *TOGAF Version 9.1*. Zaltbommel: Van Haren Publishing.

How to Build the Business Case for Enterprise Architecture. (2011, November 3). *How to Build the Business Case for Enterprise Architecture*. Retrieved July 13, 2014, from http://resources.troux.com/blog/bid/76621/How-to-Build-the-Business-Case-for-Enterprise-Architecture

Information Security and Data Breach Report. (March 2014). Navigant. Retrieved Sept 213, 2014 from http://www.navigant.com/~/media/WWW/Site/Insights/Disputes%20Investigation

Dan Lyon, danlyon@mac.com

s/Data%20Breach%20Annual%202013_Final%20Version_March%202014%20issue%202.ashx

Landau, S. (2014). Security and Privacy: Facing Ethical Choices. *IEEE Security & Privacy, 12*(4), 3-6.

Maass, E., & McNair, P. (2009). *Applying Design for Six Sigma to Software and Hardware Systems.* Prentice Hall.

ModelRisk Help. (n.d.). Retrieved September 6, 2014, from http://www.vosesoftware.com/vosesoftware/ModelRiskHelp/index.htm#Distributions/Continuous_distributions/PERT_distribution.htm

Nema, W. (2006, January). Planning Technical Architecture. *Planning Technical Architecture*. Retrieved July 13, 2014, from http://msdn.microsoft.com/en-us/library/bb245660.aspx

Nord, Robert., Barbacci, Mario., Clements, Paul., Kazman, Rick., Klein, Mark., O'Brien, Liam., & Tomayko, James. (2003). *Integrating the Architecture Tradeoff Analysis Method (ATAM) with the Cost Benefit Analysis Method (CBAM)* (CMU/SEI-2003-TN-038). Retrieved September 13, 2014, from the Software Engineering Institute, Carnegie Mellon University website: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=6557

Schneier, B. (2008, September 1). Security ROI. *Blog*. Retrieved July 13, 2014, from https://www.schneier.com/blog/archives/2008/09/security_roi_1.html

2014 Cost of Data Breach Study: Global Analysis. (2014, May 5). Ponemon Institute. Retrieved June 14, 2014 from http://www-935.ibm.com/services/us/en/it-services/security-services/cost-of-data-breach/

2013 Cost of Data Breach Study: Global Analysis. (2013, May 28). Ponemon Institute. Retrieved June 14, 2014 from http://www.ponemon.org/local/upload/file/2013%20Report%20GLOBAL%20CODB%20FINAL%205-2.pdf

2011 Cost of Data Breach Study: United States. (2012). Ponemon Institute. Retrieved Sept 14, 2014 from http://www.ponemon.org/local/upload/file/2011_US_CODB_FINAL_5.pdf

Dan Lyon, danlyon@mac.com