**John Leach Information Security Ltd**

# TBSE

## An Engineering Approach to the Design of Accurate and Reliable Security Systems

11th November 2003

An article for publication

Written by
Dr John Leach
John Leach Information Security Ltd

# TBSE – An Engineering Approach to the Design of Accurate and Reliable Security Systems

*For many years, the IT Security industry has been trying to devise a way to quantify risk and the benefits provided by security countermeasures in a form meaningful to senior business management. TBSE (Threat-Based Security Engineering) is a fresh approach to modelling and forecasting information security risk. TBSE takes a non-deterministic approach to modelling how security threats interact with countermeasures enabling quantitative forecasts of the likelihood and characteristics of security incidents as a direct function of the security measures employed. Preliminary results are encouraging and there appears to be no reason why the TBSE techniques could not be applied to a wide range of threats and countermeasures. Assuming they can, these techniques could become the foundation for a greatly-needed disciplined engineering approach to the design of accurate and reliable security systems. Amongst the many other benefits, this would give senior business management the much sought after tools with which to oversee and direct corporate security expenditures. This article describes the TBSE approach and what it can do.*

Information Security practitioners have long sought, without success, methods which would enable them to forecast in a reliable and measurable way the effects of different security measures on the level and nature of security incidents experienced. Without such methods, practitioners do not have satisfactory ways to quantify security risk and security benefits in a form which business management finds helpful, or to design security systems which can be proven to address accurately the needs of the business and provide a reliable level of risk reduction.

As a consequence, senior business management is not able to:

- Evaluate whether a given security programme will provide adequate protection to the business;
- Measure the outcomes or benefits attributable to any security programme, or evaluate the return on investment achieved by expenditure on a given security programme;

- Demonstrate to stakeholders, including shareholders and regulators, that the senior management of the company has an appropriate and cost-efficient security programme in place given the security needs of the business.

These are not small shortcomings. TBSE has been created as a way to overcome the central problem behind these shortcomings, the problem of modelling and forecasting security outcomes.

## THE PROBLEM WITH DETERMINISTIC MODELLING APPROACHES

Much effort has been spent over many years attempting to quantify risk. The overwhelming majority of these attempts have tried to deal with risk in a deterministic way. Security incidents are treated as being essentially deterministic, thought of as being caused by well-defined predictable and repeatable interactions and events. Put a certain set of circumstances together, including the full details of a specific threat and specific countermeasures, and the outcome could be fully determined as either certain success or certain defeat.

The belief underpinning a deterministic approach is that if one knew everything which one needed to know about the threat and the security measures employed, one could predict exactly what the various outcomes would be and the attendant likelihood of each outcome. However, whether or not this belief is sound in principle, it is of no use in practice. Much of what, within a deterministic model, one would need to know in order to determine the success of an attack simply cannot be known. The form in which a security incident occurs is determined by a huge number of details many of which cannot be known. For example, with a hacking attack, one cannot know the motivation or state of mind that day of the hacker, the suitability of their individual experience and skill at addressing the particular challenge the attacked security defences put before them, the amount of time they will keep going before they get bored, and so forth.

The problem with deterministic approaches comes from the fact that each security incident is essentially unique. An incident can be seen as the particular outcome of a particular set of circumstances which just happened to come together in a particular way at the time the incident occurred. It is always possible to substantiate that an incident which happened to one organisation would not happen like that to another organisation, because the details of how

they run their IT service, or the controls they have in place, or the strength of their security culture, or the way their staff are trained, will be different. It is this uniqueness of incidents which ensures that any deterministic approach will inevitably fail due to the impossibility of ever being able to capture sufficient data to support the analyses required.

An analogy will serve to substantiate this claim.

Consider medieval siege warfare and an attacking army firing a continuous shower of arrows against a castle's stone wall. If an arrow strikes solid stone, it will fall uselessly to the floor. If it just happens to find an arrow loop, the slit in the stone wall through which the defenders fire arrows out, the arrow will go through the wall and will have a good chance of wounding or killing a defending soldier. If one knew everything there was to know about the lie of the land, the weather, the time of day, the light, the bowman, his health and his state of mind, the castle and its defenders, maybe one would be able to determine the precise flight of an arrow and determine whether that arrow would find an occupied arrow slit or hit solid stone. However, such a deterministic analysis is clearly completely impractical as so many of the details essential to the calculation cannot ever be known.

Consider now the siege from the point of view of the defender firing arrows out through an arrow loop. The defender doesn't need to know which arrow from which bowman is going to be the one which finds his arrow loop. He is interested in knowing just what the probability is of any arrow from any bowman finding his arrow loop. Hence, he doesn't need to know every detail about every arrow fired. He needs to know only the density of arrows in the arrow storm and the size of the hole his arrow loop makes in the stone wall in order to calculate the chance of any arrow striking his arrow loop. With the right description of the shower of arrows, a description which requires just a fraction of the total amount of information which would be required for even a single deterministic analysis, the defender could calculate the probability of an arrow coming through his arrow loop and wounding him.

Quickly consider a second analogy, water pressure against the hull of a submarine. An engineer can work out the rate at which water will leak through the hull by knowing just the external water pressure and how watertight the hull is. The engineer doesn't need to know which water molecules will be the ones to

get through, just the rate at which water will get through as a function of the water pressure against the hull and how watertight the hull has been made.

Each security incident is essentially unique, just as are the individual arrows which make it through the arrow loops or the individual water molecules which need to be pumped out by the submarine's bilge pumps. Provided the engineer does not need to know exactly which instance of the threat will be the one which defeats the defences, he does not need to try the clearly impractical task of conducting a deterministic analysis. He can model the threat and the fraction of the threat that will get through his defences with a tractable non-deterministic analysis based on a very much simpler description of the threat. A non-deterministic analysis of the security threat interacting with security countermeasures will allow a security engineer to calculate the rate of security incidents as a function of the effectiveness of the security measures in much the same way that a castle engineer could calculate the likelihood of an arrow passing through an arrow loop as a function of the surface area of the arrow loop or a submarine engineer could calculate the rate of leakage of water as a function of how watertight the hull had been made.

## THE TBSE APPROACH

TBSE takes a non-deterministic approach to modelling risk. It models the interactions between threats and security measures in such a way that the likelihood and characteristics of security incidents can be forecast as a function of the security measures employed. This is, in outline, how TBSE works.

Security measures are classified into three classes. The first class brings together those measures which resist a threat and work to prevent attacks from being successful. Anti-virus, patching, firewalls, access control, are all examples of this type. The second class brings together those measures which do nothing to stop an attack being successful, but which work instead to make the attack's effects less severe. Developing an incident response capability is an example of this type. An incident fire-fighting team doesn't come into play until after an incident has occurred, so its primary goal is not prevention but rather to arrest the attack and contain the severity of its effects. The third class contains those security measures which do nothing to make the incident less likely or less severe but which work instead to reduce the impact, the resultant business

pain, caused by the incident. Insurance and fall-back processes are examples of this type.

TBSE is careful to distinguish between the severity and the impact of an incident. Severity is a measure of the intrinsic magnitude of the disruption caused by an incident. Impact, which depends on severity as well as on other factors, is a measure of the business pain which the induced disruption causes, i.e. the actual loss of revenue or value caused to the business by that disruption. Consider the example of an outage. The severity of an outage is measured by its duration. It is clear that the impact of an outage will vary according to the severity (duration) of the outage. In addition, the impact of an outage of a given severity will vary according to the sensitivity of the business to the processes affected by the outage. A two hour outage of a bank's trading floor will doubtless have a larger impact on the bank's business than a two-hour outage of one of its rural branches.

The likelihood of an incident occurring is controlled by the interaction between the threat and the first of the three classes of security measures. Within TBSE, this first class of security measures is referred to as "Interdictive" security measures as these work to resist the success of the threat. Interdictive interactions are modelled, in a non-deterministic way, as if they were a contest between a threat with a certain ability to penetrate defences and a countermeasure with a certain ability to resist that penetration. Clearly, this modelling technique will work only if it is possible to quantify the penetration ability of a threat and the resistance ability of a countermeasure and then to derive the function which describes how the two interact. It turns out that this is quite simple to do within the TBSE approach.

A sketched-out example will help to illustrate how this is done.

## EXAMPLE OF APPLYING THE TBSE MODELLING TECHNIQUE

Consider the simple scenario of a desktop PC on a broadband connection to the Internet protected against e-mail-borne viruses by Anti-Virus (AV) software. The ability of the AV software to prevent any virus exposure (the threat) causing an infection of the desktop (an incident) is determined by the promptness with which the user updates their desktop's AV signature files. (This presumes, which is a fair approximation in today's mature AV marketplace,

that an AV product will have a 100% rate of success at resisting any virus which is recorded in its signature file.)

In this case, the variable which describes the resistive strength of this particular Interdictive countermeasure is the number of hours between successive signature file updates. Let's call this variable β. The corresponding variable which describes the virus threat's penetrability is then the age of the virus in hours since it was first released. Let's call this variable α. The resistance function, R(α,β), which describes whether a virus of age α hours defeats an AV countermeasure updated every β hours is then a simple function of α and β and a third parameter, $T_{update}$, which describes how long, typically, it takes the AV vendor to make a new virus' signature available for users to download. It might look something like the curve shown in Figure 1 below.

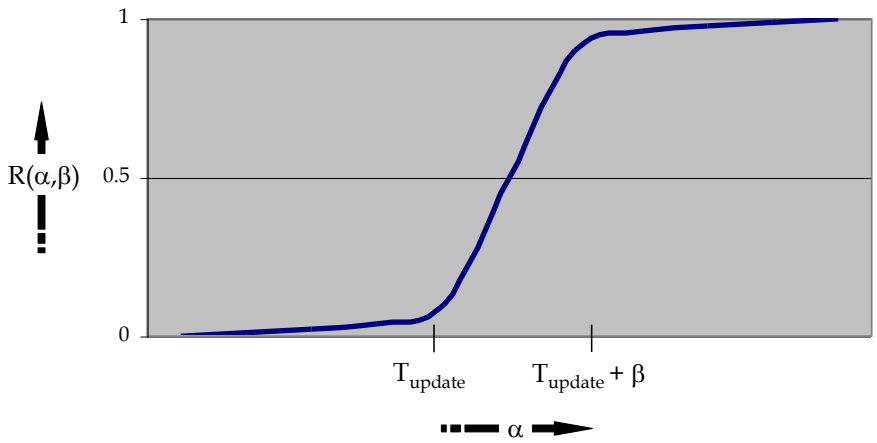


**Figure 1: The Resistance Function, R(α,β), as a function of α for a given $T_{update}$ and β**

The probability of virus infection at the desktop is determined from the resistance function, R(α,β), and the virus threat number density, n(α), the proportion of e-mail messages carrying viruses of age α. That number density could be measured readily by managed e-mail service providers. They could record over a period of time, say a month, the average proportion of e-mail messages carrying a virus as a function of that virus' age in hours. A virus threat number density might then look like the example shown in Figure 2.

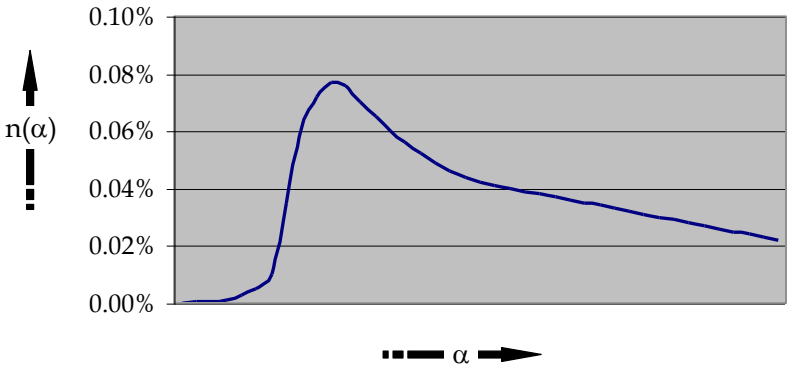


**Figure 2: A possible number density for the virus threat, n(α), as a function of α**

The probability per Internet e-mail message received of the desktop getting a virus infection can be calculated as a function of the update period β, where it should be expected that the probability of virus infection will grow with the update period β. This is indeed what the results predict, as is shown in Figure 3 for the above example data.

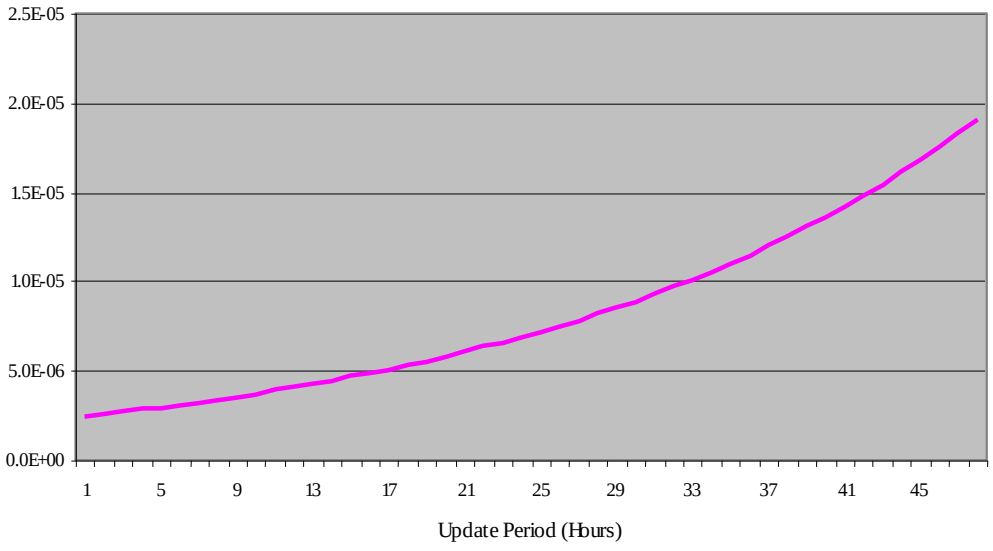


**Figure 3: The probability of virus infection per Internet e-mail received as a function of the user's signature file update period, β**

A user can then decide what probability of infection they are willing to tolerate for a given number of Internet e-mails received and determine how frequently they need to update their AV signature files in order to achieve that desired level of protection. This is shown, for the example data, in Figure 4.

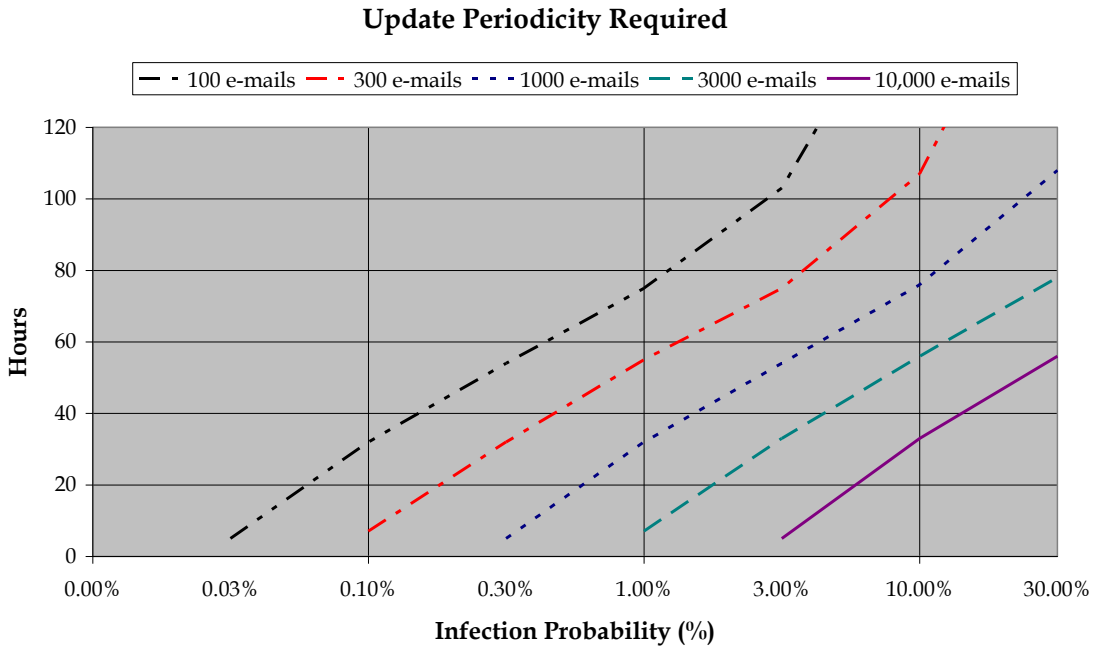


**Figure 4: AV update periodicity in hours required to achieve a given level of protection for a given volume of e-mails**

Figure 4 can be read as follows[1]. The "300 e-mails" line shows that, for the user to have no more than a 1% probability of becoming infected with an e-mail virus after receiving 300 Internet e-mails, they need to update their AV signature file within 55 hours of their AV vendor posting virus signature updates. For that probability to fall to 0.3%, the user's updates need to happen within 32 hours. For it to fall to 0.1%, the user's updates need to happen within 7 hours.

The "10,000 e-mails" line shows that, for the user to have no more than a 10% probability of becoming infected with an e-mail virus after receiving 10,000 Internet e-mails, they need to update their AV signature file within 33 hours of

[1] Do bear in mind that the results given here are for the example data which might not reflect accurately the actual number density of the virus threat on the Internet.

their AV vendor posting signature updates. For that probability to fall to 3%, the user's updates need to happen within 5 hours. The user cannot achieve a lower than 2% probability of becoming infected from 10,000 e-mails because, in this example, the AV vendor updates are not posted promptly enough. This example has the AV vendor updates posted on average 20 hours after the virus is released into the wild. This shows that, no matter how promptly the user applies available AV signature file updates, there is always a chance they will be infected before their vendor makes the virus' signature available for detection. This is an expected result which TBSE can quantify.

One can also calculate how trends in the nature of the virus threat would change the curves in the above figures. For example, if there were a trend for e-mail viruses to propagate more rapidly, perhaps if e-mail viruses increasingly carried with them their own mail servers so that they could propagate without user action, the peak of the measured $n(\alpha)$ curve in Figure 2 would move to the left, to lower values of $\alpha$. This would cause the curve in Figure 3 to move up the chart to show higher probabilities of infection at all update periodicities across the range. All the lines in Figure 4 would then move to the right. As a consequence, if a user wanted to maintain their level of AV protection despite this trend, they would need to apply AV signature updates more frequently. This is the result we would expect.

## OTHER MORE COMPLEX SCENARIOS

The above worked AV example serves to illustrate the type of results which TBSE can generate. For that purpose, the example needed to be a simple one. However, this is not to suggest that TBSE applies only to simple scenarios.

In the example, it was implicitly assumed that the eventual infection of the desktop would not change the subsequent number density of the virus threat. This is clearly a suitable approximation for the Internet virus threat for which the magnitude of the threat felt at any time is the aggregated effect of many thousands of on-line infected hosts around the world. It is clearly not a valid assumption for the internal virus threat where, once one desktop on an office LAN has become infected, that desktop then contributes to the threat environment for all the other desktops on the LAN. In these situations, the threat modelling needs to model the threat coming simultaneously from more than

one threat environment, and to model the evolution of the threat number density function over time. The TBSE technique models this quite naturally.

Multiple security measures, each contributing resistance to a given threat, can be included within the analysis by incorporating the resistance functions for each countermeasure. For example, if we take the above virus example and convert it into a worm example, a worm can be blocked in either of two ways: by its signature being in the AV signature files or by software patching to remove the software vulnerability the worm exploits. The two Interdictive measures, AV software and software patching, can be treated as working in series, i.e. the threat has to dominate both the AV countermeasure and the patching countermeasure to be successful. The TBSE technique models this quite naturally, too.

Early results show that this modelling technique can be applied effectively to a variety of Interdictive security measures, including anti-virus, patching and firewalls. The TBSE technique can also be applied to modelling internal threats where the threat comes primarily from the different types of inappropriate behaviour of staff. TBSE also models the interactions between the threat number density and the other two classes of security measure, those which make incidents less severe and those which make the impact less painful. The modelling of these interactions is conducted in a different way (though still non-deterministically) from the modelling of the Interdictive interactions, reflecting the different nature of the interactions.

There appears to be no reason why this type of analysis can not be extended to accommodate a wide variety of countermeasures, permitting the modelling and forecasting technique to be applied to a corporate IT infrastructure where host computers might be exposed to a combination of threat environments, where threats are resisted by a combination of security measures, including technical and management measures, and where the threat number density varies with time. Some of the countermeasure resistance functions will initially be harder to derive than others. In such situations, the resistance function might need to be approximated until, with the analysis of gathered data, it can be estimated more accurately and then modelled analytically.

## WHAT TBSE CAN OFFER

As was noted at the beginning of this article, the Information Security industry has in the past been unable to devise adequate methods for quantifying risk and forecasting security outcomes. TBSE is a process and set of techniques for forecasting the quantified security outcomes resulting from the use of various amounts of different security measures. It offers us the ability to design and build security measures or security programmes which are:

- Accurate – by enabling the designer to tailor and scale the design according to known and measured threat environments;
- Reliable – by enabling the designer to base the security design on the known effectiveness of each security measure at resisting, mitigating or alleviating relevant threats, thereby reducing the risk of the security design being either under- or over-engineered;
- Provable – by having the security design meet objective and measurable security targets derived from a suitable analysis of the business need for protection;
- Effective and cost-efficient – by enabling the designer to optimise the security design in order to minimise, in accordance with stated business priorities, the cost of the design or the impairment of other operational requirements such as performance, ease of use, reliability.

The return on investment expected from a security design can be calculated by costing the design and comparing that with the forecast benefits achievable as a direct result of the proposed security expenditure. Calculations can be verified independently to give management assurance. Expenditure on security measures can be justified in terms which relate to the ensuing business benefit, making it easier for management to authorise appropriate security expenditures. The effectiveness of an implemented security programme can be measured and its success or failure at satisfying the security targets set for it can be established objectively according to agreed criteria.

TBSE will bring improved transparency to the way in which Information Security supports the business, and senior business management will be able to strengthen its oversight and supervision of its Information Security arrangements. Management will then be able to demonstrate to stakeholders (includ-

ing shareholders and regulators) that they have good governance of their Information Security risks and an appropriate and cost-effective security programme in place given the needs of the business.

In a wider context, TBSE can facilitate the development of a standard framework for quantifying and calibrating security solutions, facilitating openness and interoperability. It can provide opportunities for service providers and product vendors to develop new services which quantify and forecast threat levels in useful terms. And it can facilitate the development of a fully functioning digital risk insurance market around products which can be reliably priced, offer effective cover and are simple to administer.

## CONCLUSION

TBSE is a fresh approach to modelling threat dynamics which offers an attractive opportunity for the Information Security industry to develop a disciplined engineering approach to the design of accurate and reliable security systems and which will allow security designers to show business management the benefits and return on investment they can achieve from their security expenditures.

Its application has already shown that meaningful results can be generated, and indicates that these techniques should be applicable to a wide range of threats and security measures in a practical way. A programme of work is being developed which will allow these early results to be built upon and the merits of the TBSE approach and techniques proven. The opportunity exists within this programme for companies to express their interest in the TBSE approach and to contribute to or collaborate in TBSE projects. Any such expressions of interest, or requests for a more detailed description of the TBSE approach, should be directed to the author.

Written by:
Dr John Leach, John Leach Information Security Limited
☎ = (+44) (0)1264 332 477; ✆ = (+44) (0)7734 311 567
✉ = john.leach@jlis.co.uk