10th INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM

THE FUTURE OF C2

Information Security Valuations: Definitions, Structure and Properties

Author: Michael E J Stubbings

Organisation: QinetiQ

Woodward Building
Malvern Technology Centre
St Andrews Road
Malvern
Worcestershire
WR14 3PS
United Kingdom

Tel: +44 (0)1684 895845
Fax: +44 (0)1684 896660

Email: mstubbings@qinetiq.com

Abstract

This paper examines the terminology and definitions associated with information security metrics. It summarises the results of an extensive literature search, and draws conclusions about some current approaches. The paper adopts for generic use the neutral term 'valuations' rather than 'metrics', because of the range of opinions on what constitutes a 'metric'. The paper proposes a structure which identifies and defines a set of classes of information security valuation, which relate to each other in a manner analogous to the relationship between Data, Information and Knowledge. The objective is to allow management decision-makers a greater awareness of the nature and provenance of information security valuations presented to them as supporting evidence for recommendations, thus permitting reasoned judgements about the weight to be attached to such evidence.

Introduction

This paper is derived from a study commissioned by the Ministry of Defence (MoD) into information security 'metrics' in general, and 'measures of effectiveness' and 'resilience metrics' in particular. The full MoD study examined how more clearly-defined valuations might contribute to decision-making in respect of MoD project procedures, and whether the Balanced Scorecard (Kaplan and Norton (1996)) was a suitable candidate approach for selecting information security valuations and communicating them to decision-makers. The following topics were not addressed in the study because they are considered in other, specialist, forums: Statistical analysis; Safety-critical systems; Cryptographic assurance; Formal mathematical modelling.

In the course of the study it became clear that the generic term 'metrics' also has precise and limited meanings. This study has therefore adopted the neutral term 'valuations' to describe the general topic of 'things about security to which we put numbers', using 'metrics' in a more specific sense. It also became clear that there exists no broadly accepted way of distinguishing between objective and subjective valuations, between those which describe the present situation and those which postulate some set of future events, and between a valuation describing a single phenomenon and one which is a composite representation of several phenomena.

In order to answer the MoD's questions it was therefore necessary first to step back and consider the definition of terms, establishing a known starting point both for the further explorations which have now taken place and for those which have yet to take place. It is that starting point which is presented in this paper.

Literature Search Summary

The following points emerged from the literature study:

a) Amongst the relatively few writers who attempt definitions (e.g. of 'metrics' or 'measures') there are some points in common, but no unanimity;

b) There appear to be two main categories of valuation: objective and subjective. Terminology as currently used does not support a clear distinction between objectively and subjectively derived values;

c) Discussion about using return on security investment (ROSI) values tends to focus on the relative priorities of different threat areas at a very coarse level of granularity, e.g. the relative benefits of investing in a firewall to guard against hacking, versus a single sign-on system;

d) ROSI discussions generally involve terms relating to risk management. It might therefore be reasonable to consider ROSI as the application of financial values to quantitative risk methods. Some writers cast doubts upon the usefulness and viability of ROSI calculations;

e) Some writers appear to consider 'metrics' (i.e. valuations) in general as an extension of quantitative risk assessment approaches;

f) There is a general acceptance of the difficulty, or even impossibility, of extending formal or semi-formal methods from rigorous knowledge of individual system components to rigorous knowledge of the whole system;

g) 'Metrics' are generally accepted as a good thing (e.g. 'one cannot manage what one cannot measure'), and it is broadly accepted that security investment has to be justified. There is, however, little discussion of actual management decisions to be informed by security valuations beyond the very high-level ones mentioned above in respect of ROSI, and little which relates to the system life-cycle;

h) While there is some discussion of valuations to support investment decisions, there is little discussion of post-investment review;

i) Most of the valuations mentioned in the identified sources relate either to incident statistics (e.g. number of detected attacks), or to aspects of an overall information security management regime (e.g. Is there a corporate security policy? Are all staff given security training?);

j) There has been little discussion of 'resilience metrics' (the term specified by MoD) in the context of information security. The subject appears to be related to Measures of Effectiveness, to the extent that it is questionable whether the term 'resilience metrics' describes a separate concept;

k) Most of the research in this area comes from universities in the USA sponsored by the US Government, primarily the Department of Defense.

Current Approaches to Terminology

'Measures' and 'Metrics'

Alger (2001) observes that 'a measure depends on counting', and 'metrics derives from the analysis of measures, and metrics contributes to the making of meaningful decisions and the identification of meaningful conclusions'. Vaughn (2001) observes that 'only when we can relate individual measures to some common point do they become metrics'. Vaughn separates comparison with this 'common point' from the comparison of individual measurements with each other, and thereby appears to be introducing the concept of a baseline.

Bodeau's (1995) approach includes a metric's effectiveness in supporting decision-making as one its components. She states that information security 'metrics' can be qualitative or quantitative, and does not separate a metric's identification, definition and collection from the use which will be made of it, noting a range of use-related qualities that 'good' metrics will display. These are discussed later in this paper.

Swanson *et al* (2002) have a broader approach, stating that 'While a case can be made for using different terms for more detailed and aggregated items, such as "metrics" and "measures", this document uses these terms interchangeably.'

Kovacich (1998) considers information security management regimes. He defines a metric as 'a standard of measurement using quantitative, statistical, or mathematical analyses', and states that 'In an InfoSec program, metric refers to the application of quantitative, statistical, or mathematical analyses to measuring InfoSec functional trends and workload.' Kovacich places individual metrics firmly in the context of their use, as does Bodeau (1995), although there are also similarities with Alger and Vaughn, in that he implies that a metric is something more than a raw measurement; it contains at least some element of analysis or calculation.

Measures of Effectiveness

The term 'Measures of Effectiveness' was specified by the MoD when commissioning this study, and its use seems to be confined to the defence community. It is, however, defined as follows in SANS (2003):

> 'Measures of Effectiveness is a probability model based on engineering concepts that allows one to approximate the impact a given action will have on an environment. In Information Warfare it is the ability to attack or defend within an Internet environment.'

McInerney and Montgomery (2003) discuss network enabled capability (NEC), noting C2 'measures of merit' to which 'measures of effectiveness' contribute. They state that measures of effectiveness focus 'on the impact of C2 within the operational context', quoting examples such as 'proportion of targets or threats destroyed', and 'proportion of tasks requiring co-ordination that are successfully executed'. These ideas are further explored in CCRP (2002).

Burrows *et al* (undated) states that Measures of Effectiveness are used as yardsticks 'to assess the demonstrated ability of a system to meet stated requirements'. They are 'typically derived from needed system characteristics'.

Dinolt (2003) has a more formal approach. His examples are:

- Is there a Security Policy?
- Is there a Mathematical Model of the Policy?
- How Transparent is the Mapping between the Textual Policy and the Mathematical Model?
- How was the "Consistency" of the Mathematical Model Shown?

From the second and third of these approaches, it would appear that a Measure of Effectiveness relates to a system's ability to deliver a required, defined, function, but not necessarily a security one. It is SANS (2003) which would appear to be out of step. The SANS Glossary does, however, have something in common with Dinolt's approach in that both are clearly related to formal numerical disciplines.

Resilience Metrics

The other term specifically mentioned by the MoD for this study is 'resilience metrics', but it was not found during the research. The term 'measure of resilience' was found, mostly related to environmental and social sciences. References to information security used the latter term as a synonym for some undefined ‘degree of resilience’.

There appears to be an overlap with the concepts of ‘survivability’ and ‘dependability’, as discussed by Knight and Sullivan (2000). They define a ‘survivable system’ as ‘one that has the ability to continue to provide service (possibly degraded or different) in a given operating environment when various events cause major damage to the system or its operating environment’. They also noted that survivability was one aspect of an overall concept called ‘dependability’.

Air traffic management (ATM) (Sharples (2002)) produced this definition: ‘Resilience is the extent to which the ATM system responds to a safety-significant event without causing more such events’, quoted from INTEGRA Metrics and Methodologies; Safety Metrics Technical Definitions Version 0.B 25/11/00.

Both of these approaches describe system behaviour under operational conditions and are therefore similar to definitions of ‘Measures of Effectiveness’ discussed above. Because of this overlap, the limited amount of usage for resilience-related terms, and the absence of any contra-indicated definitions, a measure or metric of (security) resilience would seem to be an example of a Measure of Effectiveness.

Return on Investment (ROI)

ROI, along with its companion term Return On Security Investment (ROSI), was not mentioned by MoD when commissioning this study. It was included because what one gets for one’s money could be seen as a measure of how effective information security investment has been.

The Department of Commerce, Government of New South Wales (2003) published a high-level guide to several approaches to security investment appraisal. It covers the following techniques: Annualised Loss Expectancy/Annual Savings model; Security Attribute Evaluation Method (SAEM); Cost-Effectiveness Analysis; Fault Tree Analysis. This document is for day-to-day use in a particular organisation, rather than presenting more abstract and academic analysis of the issues involved.

The ROI/ROSI valuations examined in this study are complex, varied, predominantly risk-based, and made up from a mixture of subjective and objective components, mostly subjective ones. This study has not examined their respective merits, and makes no comment upon them, beyond the observation that the provenance of the varied components of an ROI/ROSI figure is not always entirely obvious.

Not everyone writing on the subject of ROSI believes such a thing to be possible, at least with the degree of precision implied by its various numerical models. Heiser (2002) notes that ‘Even if ROSI was a legitimate model for security expenditures, we just don’t have the data to calculate it…Can you quantify the reduction in your firm’s future revenue streams if your e-commerce server is prominently hacked?’ Cresson Wood and Parker (2004) conclude their article by saying ‘We agree with the increasing number of experienced information security specialists who believe that ROI and similar financial decision-making methods do not apply to information security’.

Valuations for Risk and Trust

The use of numbers to represent levels of risk is frequently described in the literature surveyed as ‘metrics’, although without definitions of what constitutes such a metric.

It is not therefore possible to discuss what 'information security metrics' means without noting the various ways in which risk and trust are set out. This section of the paper summarises several approaches, including those which are commonly used in the United Kingdom public service.

Scholtz (2002) extends the discussion of risk to include an assessment of trust, complementing the concept of domains. Scholtz does not define domains, but discusses them as business-related groupings with a common security requirement. One domain might engage in business transactions with another domain and therefore have trust requirements of that other domain. Scholtz offers a 'Trust Measurement' to set out this requirement: 5 values on a scale of 0-4, with descriptions of what each value might imply. This approach has resonances with the UK guidance enshrined in HMG Infosec Standard No 3 Connecting Business Domains (Cabinet Office (2001)), where Levels of Interconnection are calculated, with a 'weighting factor' assigned to indicate the amount of confidence a domain can have in its proposed partner. The domain concept has been extensively developed and documented by QinetiQ (2001, 2003) as part of its support to MoD, and it is now MoD's preferred approach to identifying security requirements. It identifies 5 categories of information security risk, labelled A-E, these being functions of a domain's protective marking, the security clearance of its users, and the type of interconnection.

One of the well-known examples of risk valuations is the output of CRAMM (CCTA Risk Assessment and Management Method): the UK Government's preferred tool for risk assessment and management. It is software-based, marketed under licence from the UK Government by Insight Consulting (Insight (undated)). Information is gathered by means of questionnaires, the responses being entered into the tool. A model is gradually built up of assets and their valuations (on a scale of 1-7 or 1-10 depending on the asset property in question). Threats are graded as Very Low, Low, Medium, High or Very High. Vulnerabilities are graded as Low, Medium or High. CRAMM then calculates an integer 'Measure of Risk' for each of the assets or asset groups, in the range 1-7.

HMG Infosec Standard No 1 (Cabinet Office (2003)) is the HMG tool for determining the security requirements for systems processing HMG information. Use of this Standard produces a number in a range up to 21.5, indicating the level of residual risk (Residual Risk Indicator – RRI) carried for each significant security barrier for each of the following: confidentiality, availability and integrity. The RRI is a function of various parameters, including protective marking (itself arguably a 'measure' for asset value), quantity of data, and the number of potential attackers. A table in the Standard sets out how an RRI can be read across into an Evaluation Assurance Level (EAL), thus establishing the degree of Common Criteria evaluation assurance considered appropriate to the security barriers in question.

Technical Assurance – Common Criteria

One of the best known uses of numeric scales in information security is the Common Criteria regime. This international initiative provides mutual recognition for Certificates identifying the degree of security assurance (the Evaluation Assurance Level, or EAL) provided by specified components of a product or system, on a scale of 0-7. An EAL number would appear therefore to be an example of an information security valuation. Further information about the Common Criteria approach can be found on the website of the UK National Technical Authority for information security, the Communications-Electronics Security Group (CESG (undated)).

CESG runs similar schemes, including SYSn assurance packages, described in the SYSn Assurance Packages Framework (CESG (2002)). Other technical assurances which could be seen as valuations include CESG's Fast Track Assurance service, the CESG Approved Products Scheme (CAPS, for commercial cryptographic products), and the USA's Federal Information Processing Standard (FIPS) 140, also for commercial cryptographic products.

Compliance and Audit - NIST Guidebook

The US National Institute of Standards and Technology (NIST, part of the US Department of Commerce) has published a Guidebook on the subject of security metrics (Swanson *et al* (2002)). The amount of US Government material on this and allied topics is considerable, and the NIST document will be taken as representative of their approach to security valuations for compliance and audit.

The document discusses the establishment, development and maintenance of an information security metrics programme, exploring typical information security roles and responsibilities within US Government organisations. It is orientated around the duties of such organisations' staff, and their obligation annually to report their information security posture to the Office of Management and Budget.

The sample values included in the document are exclusively concerned with the security management regime. Examples include:

a) Percentage of systems that have the costs of their security controls integrated into the life cycle of the system;

b) The average time elapsed between vulnerability or weakness discovery and implementation of corrective action;

c) Percentage of information systems libraries that log the deposits and withdrawals of tapes.

Systems Security Engineering Capability Maturity Model (SSE-CMM)

The Systems Security Engineering Capability Maturity Model (SSE-CMM (undated)) is a collaborative effort by US and Canadian defence and intelligence authorities, along with a number of industry partners. SSE-CMM is a model for assessing the effectiveness of the development environment within which systems security components are built and tested. The model has achieved ISO status (ISO/IEC 21827:2002). The body charged with further development of SSE-CMM is the International Systems Security Engineering Association (ISSEA). It maintains a Metrics Working Group, currently conducting further research into information security metrics. The model allows assessment of an organisation's project and security engineering functions against five gradually increasing levels of capability: Performed Informally; Planned and Tracked; Well-Defined; Quantitatively Controlled; and Continuously Improving. There are similarities to BS7799 in that it allows external validation of declared aspects of an organisation's security management regime, in this case, those relating to software development.

The SSE-CMM approach follows Bodeau (1995) in several significant respects. For SSE-CMM, a security metric can be objective or subjective, quantitative or qualitative. Both present valuations in the context of their usefulness for informing key decisions. SSE-CMM discusses metrics as being either Process Metrics or Security Metrics. The former relates to an organisation's ability to demonstrate the maturity of its security processes, and the latter to the results of its security processes. Bodeau discussed metrics as being either Process or Product in focus, although she introduced the extra concept of the Requirements security metric.

BS7799

There is a formal audit process to support Part 2 of British Standard 7799 – Information Security Management Systems (BSI (2002)). Part 1 of BS7799 is a Code of Practice, also published as ISO/IEC17799 (BSI(2000)). Part 2 (not yet an ISO Standard) is the Standard against which certification audits can be carried out in accordance with procedures established by the United Kingdom Accreditation Service (UKAS). BS7799 is about the Information Security Management System (ISMS) – the regime which selects, implements, maintains, monitors and reviews information security measures. It does not specify any particular valuations, although audit practice indicates that comprehensive ISMS record-keeping is essential to a successful certification audit. Such records could be regarded as measurements for later aggregation into ISMS performance metrics. The success or otherwise of a BS7799 audit – whether for formal certification or not – could itself be regarded as an information security valuation with a binary value range.

Management Statistics

Kovacich (1998) also sets information security valuations firmly in the context of an information security management regime. He advocates the preparation of graphs to demonstrate values such as:

- Total average processing time for applications for security approval for new systems over time;
- Average number of attendees at infosec awareness briefings over time;
- Percentage of user population having user privileges revoked over time.

As can be seen from the examples above, these valuations are firmly orientated around the cost-effectiveness and resourcing issues associated with the information management regime itself, rather than those relating to individual counter-measures.

Observations on Findings

What Was Not Found

There is very little discussion of the difference between numbers based upon measurements in the physical world and those based upon estimates and subjective assessments. This study suggests that these two categories represent two very different concepts, which should not be confused with each other.

While quantified risks and return on investment have been discussed in the context of organisational strategy decisions, there is little discussion about how information security valuations support decision points in project and system life-cycles.

Much has been written on the subject of ROSI. Nothing has, however, been found on the subject of measuring ROSI actually achieved, nor on how such measurements (if they are possible) should be used to inform future decisions or to curtail nugatory expenditure. Cresson Wood and Parker (2004) make this point when they state ‘As far as the authors know, there has never been a scientifically-based study examining the accuracy, validity, or effectiveness of ROI or any other financial analysis tool as applied to information security projects’.

Much of the writing about valuations refers to figures extracted from government or commercial information security surveys. Such figures appear to be taken at face value, with little discussion of what the numbers actually describe. For example, the literature search did not reveal much questioning of how representative a survey might be in respect of the total population set of the categories being described.

The literature reviewed has largely assumed decisions based upon the comparison of like valuations. Information security managers and system designers do have, however, to consider how to make decisions such as whether to choose a procedural counter-measure to handle a particular risk rather than a technical counter-measure. Such decisions might be based upon the performance of steps already taken, but whether information security valuations could clarify this topic remains undiscussed.

The concept of 'defence in depth' has been extensively discussed and applied over the years. At present, however, there appears to have been little exploration of how information security valuations might illuminate the working of combinations of probably disparate counter-measures against specific risks.

Vaughn (2001) suggests that 'a 100% predictive measure for assurance of software intensive systems' may not be found. The possibility that a way of measuring the actual (as opposed to predicted) performance of combinations of software components might be found does not appear to have been pursued. The concept of measures of effectiveness seems relevant in this context.

Objective and Subjective Valuations

This study is intended to produce, among other things, suitable definitions of terms for information security 'metrics'-related concepts. The literature search has shown that 'metrics' fall into two categories: objective and subjective. The first question to be asked therefore is whether we need to separate these two concepts.

The application of numbers to a question, particularly when associated with formulae for their calculation or evaluation, lends credibility to the properties thus described. A value, for example of '4.2' (to select a number at random) gives an impression of precision and reliability. If it were based upon an attack impact cost derived from industry surveys, multiplied by a consultant's estimate of the current threat level (e.g. a value for 'medium') for that sort of attack, further modified by the cost of a particular proposed counter-measure, what would '4.2' actually mean? If, however, '4.2' were the average number of occasions per day on which users forgot their password out of a user community of 25 people, that might indicate something very definite about the quality of user training, and the suitability of the password regime.

This is not to denigrate the use of subjective assessments or the use of forecasts. It is simply to point out that when presenting numbers as a basis for taking security investment decisions, knowing whether a number's provenance is objective or subjective, whether it describes the past or estimates the future, allows managers more readily to understand its implications.

This paper will proceed on the assumption that decision-making clarity would be enhanced by differentiating between these two categories. This in no way demeans the value of subjective assessments, forecasts, numbering or classification systems, or calculations derived from them. It is simply a question of having words available which describe concepts unambiguously. This study asserts that objective and subjective assessments are different techniques for supporting decision-making, and decision makers should know which they are dealing with.

Valuations and the Use Made of Them

Some writers have concluded that the use made of a value is part of its definition. This raises interesting questions such as if a valuation has two uses, whether it is two valuations. Because of the lack of satisfactory answers to these questions, this study will not follow that route. This is not to down-play the importance of a valuation's usefulness – merely to make it an attribute than a prerequisite for a valuation's existence. If there were no obvious use for a particular valuation, this attribute would have a null 'value'. The attribute could change in 'value' over time as different uses appeared and disappeared.

Bodeau's appreciation of the importance of a valuation's use will not, however, be set aside. She considered the quality, or fitness for purpose, of individual valuations, and presented the use and quality of a valuation as an essential part of the valuation in question (what will be described in the Appendix definitions as a 'generic property'). Bodeau's criteria for fitness for purpose have been included in the generic properties and associated information proposed in Appendices 1 and 2.

A Proposed Approach

Measures and Metrics

This study will follow Alger's lead by defining a measure or measurement as a single objective valuation, and a metric as a value derived in some way from two or more measurements. Appendix 1 contains more formal proposed definitions for these terms, with proposed property definitions in Appendix 2.

The following examples will illustrate these definitions.

a) Certain types of events logged in a system audit trail might be defined as potential attack indicators. The number of such events detected per day would be a measurement. A related metric would be changes in the number detected per day, perhaps cross-referenced with some other event, such as the introduction of a new website, or the adoption of a new system patching regime. On a larger scale, the UK Department of Trade and Industry (DTI) information security breaches survey would be a source of metrics such as the number of reported attacks. These figures would qualify as metrics because they have emerged from a consolidation and analysis process.

b) The length of time between the arrival of a new-entrant in an organisation and their receiving a security briefing is a measurement. A matching metric would arise when such measurements are compared with the organisation's stated policy on the provision of such briefings, or compared with the length of time before new-entrants are given access to the organisation's sensitive information.

c) A new software-based system is being developed. The number of individual components which have successfully passed an independent review of their fitness for purpose is a measurement, as is whether an individual component has received such a review or not. A matching set of metrics might be a correlation of these measurements with the experience and qualifications of the reviewers, or with the number of reviews carried out per reviewer per week.

The approach in this study will set aside Vaughn's inference (Vaughn (2001)) that a predefined baseline is an essential component of a metric. This is because it is conceivable that some measurements might be gathered, consolidated, combined with other types of measurement and then presented without any form of baseline. For example, failed login attempts might be registered, and correlated with time of day and day of the week. This would impart more information about failed logins than the raw measurements would have done, but without having a baseline present.

Measures of Effectiveness

This study will follow the defence-orientated focus for Measures of Effectiveness discussed above, defining such a valuation as a comparison between an actual performance and a target performance. A more formal proposed definition is set out in Appendix 1, with property definitions in Appendix 2.

There is an overlap here with the 'common point' mentioned by Vaughn (2001). By defining a metric as including a baseline or 'common point' for comparison, it might seem that he is describing a measure of effectiveness. But a measure of effectiveness contains a target, which need not be the same thing as a baseline. A baseline might be the point at which some action is triggered such as calling out an incident response team. Such a baseline would not be a target. If, however, a baseline denoted a minimum acceptable level of performance, then it could be described as a target. It is doubtful, however, whether in operational use the distinction would have any practical significance.

A measure of effectiveness is truly a 'measure' rather than a 'metric', because it contains only one measurement – the actual performance. The other component is the target, which describes not an actual phenomenon, but a future objective. A measure of effectiveness could become a component part of a metric.

There is, however, a slight complication. An organisation might set a target for a high-level aggregation of system elements. The single 'measure of effectiveness' might then itself be an aggregation of lower-level physical observations. Under these circumstances, would it therefore be a 'metric of effectiveness'? According to the definitions set out in this study, the answer to that question is yes. But attempting to replace custom and practice with what would inevitably seem like pedantry would not add to the clarity of decision-making.

The established usage (as illustrated by CCRP (2002)) covers a range of assessments including subjective ones. A measure of effectiveness should therefore be regarded as potentially objective or subjective in nature, although some accompanying text is advisable to explain what it is and where it comes from.

The definition given above allows for a wide range of examples. These include:

a) The formal methods approach set out by Dinolt (2003);

b) Structured and defined subjective value judgements such as perceived performance for security responsibilities in an individual person's annual performance appraisal;

c) The performance of high-level aggregations of system elements, such as the ability of an entire communications suite to deliver messages to the intended recipients within a specified period of time.

Resilience Metrics

As noted earlier, the term 'resilience metrics' does not appear to refer to a separate concept despite its interest to MoD. The term 'resilience measurement' has not been observed. In the absence of any apparently established usage, this study proposes a similar distinction between measurement and metrics in the context of resilience as it has done when those terms are used alone; noting the overlap with 'Measures of Effectiveness'. A more formal definition for resilience measurements and metrics is proposed in Appendix 1, with property definitions in Appendix 2.

Risk Indicators

Return on (security) investment would seem to be a special case of risk metrics, with values expressed using a monetary scale. These are calculated from estimates of asset value, business impact, threat likelihood and other parameters, much as tools like CRAMM do for their (non-financial) risk calculations.

If, however, a metric is drawn from two or more 'measurements', then is a CRAMM number in the range 1-7, or an RRI number from HMG Infosec Standard No 1 truly drawn from 'measurements'? Do their component measurements each represent 'A value representing a single instance of a defined and observed information security phenomenon'? The answer would appear to be 'no', which means that these aggregate or calculated values cannot, in these terms, be described as metrics.

This is not to say that numerical or monetary values cannot usefully be assigned to risks or to assessments of return on investment. Those topics are, in any case, not the subject of this study. They are, however, frequently described as 'metrics'. This study suggests that this is at best unwise, as the common use of the words 'measurement' and 'metrics' implies a degree of objectivity, precision, repeatability and reliability which risk and return on investment are not in general strong enough to carry. This is because calculations of risk or return on investment are based upon estimates and forecasts. One cannot 'measure' the future.

A risk valuation is generally a single value derived, often algorithmically, from several separately assessed components (e.g. asset value, threat or probability, impact). It is generally, however, regarded as being descriptive of a single unwanted situation or event. Despite its composite and derived nature, this paper will therefore treat it as an individual discrete value describing that one situation.

This study suggests that it would be preferable not to apply the words 'measurement' and 'metric' to these calculations. They are estimates, or are derived from estimates. A number calculated from an estimate remains itself an estimate. Useful, even invaluable, it may be, but a metric or a measurement it is not. The word 'indicator' would be better for numerical scales used to denote levels of risk, as proposed in Appendix 1.

Metrics for Audit and Compliance

Compliance models are perhaps more complex than technical assurance valuations because they deal more explicitly with human behaviour, and human beings are generally more complicated than machines. SSE-CMM, BS7799 and US Federal

Government audits are highly structured ways of determining the performance of a system's developmental or operational environment, including its people.

These audit approaches are composites of lower-level measurements and metrics, and their results can thus be categorised as metrics. For example, detailed security record-keeping and statistical analysis will inform the day-to-day monitoring and adjusting of an information security management regime, and can be checked by a BS7799 audit or US Federal Government audit. SSE-CMM is similar in that it measures an engineering environment. If, however, there were a stated target result for the audit, then the actual result would become a measure of effectiveness.

From this it would seem that the concepts of measurement, metric and measure of effectiveness can be recursive. The latter can be derived from various metrics, themselves derived from measures. All three terms can properly be applied in the context of approaches to audit and technical or organisational compliance with stated targets. This study does not therefore propose terms specifically for the purposes of auditing.

Other Information Security Valuations

There are several concepts described in the source literature as 'metrics', but assessed here as something being rather different. This study proposes that the following terms and concepts be employed for them. Vaughn (2001) mentions indicators and predictors, and these terms have been re-used here. More formal definitions are proposed in Appendix 1, with proposed property definitions in Appendix 2.

a) Risk Indicators
Single values selected from a numerical or other structured classification system, describing the nature and level of a particular risk.

b) Predictors
Single values indicating some aspect of expected future behaviour, for example, the threat indices used by the Security Attribute Evaluation Method (SAEM – Butler (2001)). When combined in some consolidated or calculated form they could be absorbed into a Forecast. They would thus have the same relationship to a Forecast that a Measure has to a Metric.

c) Indicators
Single values in some numeric or other structured classification system, giving a subjectively derived description of an information security phenomenon. An example could be the grading given to the security performance of an individual member of staff. An Indicator would have the same relationship to an Assessment that a Measure has to a Metric.

d) Forecasts
Consolidated or calculated values indicating some postulated state of future measurable events. They would be calculated from Predictors or from a combination of Predictors and Measures / Metrics. An example would be the expected level of system attacks at some point in the future. Forecasts would be essential components in any presentation of expected Return on Security Investment.

e) Assessment
Informed opinions derived from two or more correlated indicators, expressed in numerical or other structured terms, or in text.

f) Targets
A special case of Predictor, describing the measurable future performance of some entity. When combined with a measurement of the actual performance, it would become a measure of effectiveness.

g) Pointers
Single values in some numeric or other structured classification system, giving the expected result of a future subjective assessment of an information security phenomenon. An example could be the level of risk expected at some specific point in the future. A Pointer would have the same relationship to a Prospect that a Measure has to a Metric.

h) Prospects
Consolidated or calculated values indicating some postulated state of future subjectively-assessed information security phenomena. An example could be the levels of risk expected at some specific point in the future.

Despite their rejection as metrics or measures, these concepts can be clearly defined, and can take their place as tools for information security practitioners and investment managers. The approach set out here allows a clear separation between those elements which are truly measured, and those which are estimated, thus giving more information to the users of such figures about their nature and provenance.

A Proposed Table of Terms

The terminology defined so far has a parallel in definitions commonly applied to the terms data and information. Smith (undated), in his paper on warfare-related decision-making offers the following definitions:

> "Data" is the raw untouched input direct from a source or sensor with no attempt made to judge its validity or accuracy.
>
> "Information" is data that have been collated to establish a relationship with other known facts.
>
> "Intelligence" is information that has been analyzed to derive the meaning and implications of the information.

Smith also notes that 'intelligence' means in this case 'knowledge of the enemy'. He also suggests that 'knowledge' could be used instead of 'intelligence'.

The terms 'data' and 'information' would appear consonant with the definitions adopted earlier in this study for 'measurement' and 'metric'. No equivalent for 'knowledge' or 'intelligence' was found in the literature studied. Several writers have noted that valuations are linked to decision-making, which would appear to be the level addressed by Smith's use of the words 'knowledge' and 'intelligence'.

We therefore appear to have three levels of terminology, broadly parallel in their definitions and use, across five categories: generic terms (addressed by Smith); objectively-based terms; subjectively-based terms; and terms relating to objective and subjective expectations for the future. This can be illustrated as follows.

GENERIC	OBJECTIVE	SUBJECTIVE/ RISK	FUTURE OBJECTIVE	FUTURE SUBJECTIVE
Data	Measure	Indicator	Predictor / Target	Pointer
Information	Metric	Assessment	Forecast	Prospect
Knowledge	Knowledge	Knowledge	Knowledge	Knowledge

Figure 1: Comparative Terminology

The first level (Data, Measure, Indicator, Predictor / Target, Pointer) is the single raw item or value, in isolation. The second level (Information, Metric, Assessment, Forecast, Prospect) is where some form of correlation takes place between two or more first level items, placing them in a more informative context.

This leaves only the term ‘knowledge’ unexplored. This study suggests that this is the point where the valuations harvested are actually used for decision-making. The following definition is suggested as being appropriate to all the categories in Figure 1.

> Knowledge is the understanding acquired by a decision-maker as a result of being supplied with metrics, information, an assessment, a forecast, a prospect, or some combination of any or all of these. It is this understanding which informs the decision subsequently made.

It will become clear from the definitions offered in the Appendices that a Forecast and a Prospect are very similar in that they are both summations of expected future events. They differ in that the accuracy of one can eventually be checked by objective measurements, and the accuracy of the other can eventually be checked by subjective judgements. A parallel point can be made about the difference, and similarity, between a predictor and a pointer. Both are values representing a single future event. The difference is that one can eventually be checked for accuracy against an objective measurement, and the other against a subjective judgement.

Approaches to Selecting Valuations

There are two main approaches to identifying the values one wants to gather and use. The first (top down) starts from requirements statements for the project, system, etc., working out what values would allow one to determine whether one had, and continued to have, a satisfactory solution to those requirements.

The second approach (bottom up) starts from what is already there, e.g what operating system log facilities, or staff recording and reporting regimes are in place. A valuation regime is then built on what is available. Birchall *et al* (2004) refer to these as ‘incumbent metrics’, noting that current reporting methods are largely bottom-up in nature.

Adopting an approach which is strictly uni-directional does have its disadvantages. It is all very well at the requirements capture stage specifying exactly what one wants to measure, but the optimum technical solution to the problem might not provide that information. If one waits until one has an operational system to see what information is available, it may not tell managers everything they want to know.

It is therefore likely that those who wish to gather and use information security valuations will use both approaches. Specifying what one wants at the requirements capture stage will be modified by an assessment of what is actually possible and of what is cost-effective. Extracting and processing what is available from an existing situation will be modified by consideration of what one wants to know, and may result in information or organisational systems being modified in order to provide this – effectively the requirements will be revised and the system modified to meet them.

Two conclusions can therefore be drawn. The first is that an effective regime for gathering and processing information security valuations is inextricably linked with the requirements capture process. The second is that such a programme has to bear in mind what is feasible with the technical and human resources available.

The usefulness of a valuation is linked to 'alignment', whereby information assurance measures are aligned with broader business and organisational objectives, and with the specific interests of the various stakeholders (customers, auditors, staff, regulators, partners etc.). This alignment is demonstrated and monitored by means of the various valuations which are the subject of this study. This subject underpins the Henley Management College information assurance report referenced as Birchall *et al* (2004). The panel of experts consulted by the writers of that report noted that the two most pressing requirements for co-ordinating information assurance (IA) practice and business strategy were: improving communications between IA and business functions; and aligning IA measures with business objectives.

It is information security / information assurance valuations which will demonstrate whether the second of these has been achieved. It is these valuations which – if well chosen - will provide a means of communication between IA and the broader business. This applies whether the business functions are those of a retail company (stock control, sales, marketing, retail outlet franchising) or whether the business functions are those typical of the MoD: logistics, secure communications, battlespace analysis and decision-making.

Summary

There is much interest in valuations as a contribution to effective information security decision-making, but current approaches render it difficult readily to identify the provenance and meaning of the valuations produced. This study suggests that the English language offers a number of common-usage terms which allow this differentiation to be made with clarity. It also suggests that there is a three-level structure common to all the categories proposed. The adoption of a structured, defined terminology of this sort would give decision makers a clearer understanding of the values put before them. It would enhance decision-making by eliminating the opportunity for inadvertent and spurious comparison between objective and subjective values, and by clarifying and separating the roles of calculation and judgement.

The requirement for a measure of effectiveness implies the presence of a target performance against which to measure the achieved performance. The presence of a target implies that the requirements capture phases of a project must address 'measurable' aspects of information security, as that is where targets are set. Opportunities for gathering other measurements (and thus metrics) will themselves depend in part on the effectiveness of the capture and statement of information security requirements. If the availability of a mechanism for capturing a particular measurement depends on it having been built into a system, then this must be identified during the requirements capture phase. The only valuations available otherwise will be those which can be gathered procedurally, or those which one is fortunate enough to acquire accidentally – e.g. because a COTS operating system happens to record a particular category of event in a system log.

There has been a considerable amount of discussion about very high-level security investment decision-making, particularly in relation to the calculation of anticipated return on security investment. Current approaches to Return on Security Investment (ROSI), however, apply high degrees of precision to what are basically risk assessment judgements made about future events. In contrast to very precise calculations about the future, there seems to have been very little discussion about how one might calculate the return on security investment actually achieved for systems which have already gone into operation. Similarly, there has been little discussion about how information security valuations might support decision-making in project and system life-cycles and in information system procurement procedures.

Future Work

Some further work has already commenced, primarily into the suitability of the proposed structure and definitions to inform and guide the selection of information security valuations, and into the potential for such valuations to be usefully included in security-related Balanced Scorecards.

Other work programmes have been identified as making further contributions to the subject of information security valuations and their use. They will of necessity explore the practical applicability and usefulness of the theoretical structures set out in this paper. These programmes are:

a) Study of the relationship between information security 'measures of effectiveness' and broader C2 Measures of Merit;

b) Study of the relationship between information security valuations and approaches to requirements capture, including the QinetiQ/MoD Domain-based security model, preferably in the context of live projects;

c) Further exploration of the concept of 'dependability' in the context of information security;

d) Further study of how information security valuations might relate to:
   - BS7799/ISO17799;
   - Systems Security Engineering Capability Maturity Model (SSE-CMM);
   - Obtaining Common Criteria assurance;
   - HMG accreditation procedures, as set out in Joint Services Publication 440 (JSP440), and in the HMG Manual of Protective Security (MPS).

Appendix 1

PROPOSED VALUATION DEFINITIONS

Assessment

A calculated, comparative or aggregate value derived from two or more indicators (q.v.), measures (q.v.) or metrics (q.v.), at least one of which is an indicator, imparting further information about the assessed phenomena

Generic Properties

Representative, Repeatable, Singular, Composite, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Inference, Discrimination, Expression

Example

Average ‘security awareness’ marking in an annual staff appraisal round for a particular department or team

Forecast

A calculated, comparative or aggregate representation of two or more elements, at least one of which is a predictor (q.v.), imparting further information about the expectations of information security phenomena yet to be observed

Generic Properties

Representative, Repeatable, Singular, Composite, Physical, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Precision, Discrimination, Expression

Example

The average number of projects expected to be allocated to each security approvals officer in a particular team

Indicator

A value representing a single instance of a judgement made about a single quality or aspect of a defined information security phenomenon, according to a specified, bounded numerical or other structured classification system

Generic Properties

Representative, Repeatable, Singular, Simple, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Inference, Discrimination, Expression

Example

A performance rating given for an individual member of staff against the 'security awareness' heading in their annual appraisal

Measure (or Measurement)

A value representing a single instance of a defined and observed information security phenomenon, according to a specified numerical scale or other objective classification system

Generic Properties

Representative, Repeatable, Singular, Simple, Physical, Attested

Specific Properties for Each Instance

Use, Owner, Source, Collection, Processing, Cost, Precision, Discrimination, Expression

Example

The number of port scans registered at a particular network node during a defined period

Measure of Effectiveness

A value, selected from a defined range of values, which describes a component's performance in discharging a defined function, compared with that component's target performance

Generic Properties

Representative, Repeatable, Plural, Attested,
Simple or Composite (established usage may involve an actual performance element which is either a measurement (q.v.) or a metric (q.v.)),
Physical (if derived from objective valuations)

Specific Properties of Each Instance

Use, Owner, Source, Collection or Production, Processing, Cost, Precision, Discrimination, Expression

Example

The number of authentication requests satisfied within a stated time by a specified authentication server under stated conditions, compared with the target rate established for that particular server under that set of conditions

Metric

A calculated, comparative or aggregate representation of two or more measurements (q.v.) which imparts further information about the observed information security phenomena

Generic Properties

Representative, Repeatable, Singular, Composite, Physical, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Precision, Discrimination, Expression

Example

The average number of port scans registered per day over the previous month

Pointer

A value representing a single instance of a judgement made about a single quality or aspect of a defined information security phenomenon which is yet to take place, according to a specified, bounded numerical or other structured classification system

Generic Properties

Representative, Repeatable, Singular, Simple, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Inference, Discrimination, Expression

Example

The expected risk severity (e.g. high, medium, low) for flood damage at a particular site, should a particular river engineering scheme be implemented

Predictor

A value representing the expected result of a future measurement (q.v.)

Generic Properties

Representative, Repeatable, Singular, Simple, Physical, Attested

Specific Properties of Each Instance

Use, Owner, Source, Collection, Production, Processing, Cost, Precision, Discrimination, Expression

Example

The expected rate of security-related helpdesk enquiries

Prospect

A calculated, comparative or aggregate value derived from two or more pointers (q.v.), predictors (q.v.) or forecasts (q.v.), at least one of which is a pointer, imparting further information about expectations of future information security phenomena

Generic Properties

Representative, Repeatable, Singular, Composite, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Inference, Discrimination, Expression

Example

An expected future decision from an information security approvals officer (e.g. a government information security accreditor)

Resilience Measure (or Measurement)

A special case of a measure of effectiveness (q.v.), describing a system's actual performance against a specific target performance, where the target performance relates to a system's ability to react to defined sets of events without impeding the operation of stated functions

Generic Properties

Representative, Repeatable, Plural, Simple, Physical, Attested

Specific Properties of Each Instance

Use, Owner, Source, Collection, Processing, Cost, Precision, Discrimination, Expression

Example

The number of network-based authentication requests capable of being handled within a stated time by a particular authentication server under stated conditions, before a stated degradation of network performance occurs, compared with the target established for minimum required authentication capacity under those conditions

Resilience Metric

A special case of a measure of effectiveness (q.v.), describing a calculated, comparative or aggregate representation of two or more resilience measurements (q.v.) which imparts further information about the system's actual performance against a single, stated target, where the target relates to a system's ability to react to defined sets of events without impeding the operation of stated functions

Generic Properties

Representative, Repeatable, Plural, Composite, Physical, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Precision, Discrimination, Expression

Example

Bottleneck identification (worst resilience result): the comparison of individual resilience measures in a processing chain to find that which most restricts system performance

Risk Indicator

A special case of an indicator (q.v.) representing a judgement made about the nature and/or severity of a defined risk applying within stated boundaries, according to a specified, bounded numerical or other structured classification system

Generic Properties

Representative, Repeatable, Singular, Simple, Attested

Specific Properties of Each Instance

Use, Owner, Source, Production, Processing, Cost, Inference, Discrimination, Expression

Example

The number assigned by a CRAMM analysis to a particular defined risk

Target

A special case of a predictor (q.v.), describing a value representing the required result of a future measurement (q.v.)

Generic Properties

Representative, Repeatable, Singular, Simple, Physical, Attested

Specific Properties of Each Instance

Use, Owner, Source, Collection, Production, Processing, Cost, Precision, Discrimination, Expression

Example

The number of security approvals staff to be employed

Appendix 2

PROPOSED PROPERTY DEFINITIONS

Generic Properties

Attested
The valuation is the result of a stated, fully defined process for collection and calculation, or in the case of anticipated measurements, will be the result of such a process, or is the result of a stated, defined set of structured judgements

Composite
The valuation is constituted from two or more components

Physical
The valuation is descriptive of an actual or anticipated physical event, or is derived from measurements of such events.

Plural
The valuation is composed of two or more separate values.

Repeatable
The same objective collection and calculation process, applied to the same phenomenon or phenomena under the same circumstances, will always have the same result.

The process for deriving a subjective present or future valuation is sufficiently well-defined and structured for similar values to be reached by independent agents under the same circumstances.

Representative
The range of possible values represents a defined range of conditions in the phenomenon being described.

Simple
The valuation has only one component.

Singular
The valuation is a single value.

Specific Properties

Collection
The procedure for collecting a measurement, e.g. locations, frequency of collection, techniques used, responsibility for collection

Cost
Cost of collection, production, processing etc. expressed as appropriate (e.g. money, resources, time, operational delay)

Discrimination
Ability of the value range to represent the entire range of possible conditions in the phenomenon it describes (as opposed to the subset described in the ‘Representative’ generic property)

Expression
Ability of the normal spoken language accurately to communicate the nature of the valuation and the number or other result associated with it

Inference
The defined set of inferences about the physical world associated with the valuation

Owner

Person or role responsible for collecting, calculating or otherwise producing the valuation, and for storing it

Precision

Level of accuracy required (e.g. to the nearest minute or to the nearest second)

Processing

Storage, recording and other post-collection or post-production processing of valuation

Production

The particular algorithms, structured judgements or other non-collection processes for producing the valuation, including frequency of production, locations, techniques, responsibility etc.

Source

The information security phenomenon, attribute or quality described by the valuation

Use

A single purpose for which a valuation is collected, stored, processed and presented. A purpose is one only of the following:

- To inform a decision, this being the selection of one or more of a finite number of defined courses of action, where a course of action can include consciously doing nothing;

- To inform an assessment, this being the application of analysis and judgement to a situation, in order to simplify and present its essence and the inferences which can be drawn from it.

Each use would be associated with the following information:

Purpose

A particular decision or assessment

User

The person or role making the decision or assessment

Schedule

The absolute or relative point in time when the valuation is required

References

Alger, John I, (2001), On Assurance, Measures, and Metrics: Definitions and Approaches; philby.ucsd.edu/~cse291_IDVA/papers/rating-position/Alger.pdf

Birchall D, Ezingeard J-N, McFadzean E, (2004), Information assurance – Strategic alignment and competitive advantage, Grist Ltd., London (report issued jointly by Henley Management College and QinetiQ Ltd)

Bodeau D J, (1995), Proceedings of the 18th National Information systems Security Conference, Baltimore, USA, October 10-13 1995, Measuring Security: What Can We Learn From Other Fields?

BSI (2000), ISO/IEC 17799:2000(E), British Standards Institution, London, ISBN 0 580 36958 7

BSI (2002), BS7799-2:2002, British Standards Institution, London, ISBN 0 580 40250 9

Burrows R, Mitchell Col T L, Ball J R, Joglekar A, Schneider Edward A Jr, (undated), Issues In Operational Test and Evaluation (OT&E) of Information Assurance Vulnerabilities, Institute for Defense Analyses, www.dodccrp.org/Proceedings/DOCS/wcd00000/wcd000b8.htm

Butler, SA, (2001), Computer Science Department, Carnegie Mellon University, Pittsburgh, USA, Security Attribute Evaluation Method: A Cost-Benefit Approach, www-2.cs.cmu.edu/~shawnb/SAEM-ICSE2002.pdf

Cabinet Office, (2001), HMG Infosec Standard No 3 (Connecting Business Domains)

Cabinet Office, (2003), HMG Infosec Standard No 1 (Residual Risk Asessment Method)

CCRP, (2002), Department of Defense Command and Control Research Program, NATO Code of Best Practice, Washington DC, ISBN 1-893723-09-7

CESG (undated) Communications-Electronics Security Group website, www.cesg.gov.uk

CESG (2002), SYSn Assurance Packages Framework, Issue 1.0, September 2002, Communications-Electronics Security Group, www.cesg.gov.uk

Cresson Wood Charles, Parker Donn B, (2004), Computer Fraud & Security (Issue 5, May 2004), Why ROI and similar tools are not advisable for evaluating the merits of security projects

Department of Commerce, Government of New South Wales, Australia, (September 2003), Return on Investment for Information Security Guideline, Issue 1.0, www.oict.nsw.au/pdf/4.4.37.ROSI.pdf

Dinolt George W, (2003), Use of Formal Methods in Assessment of IA Properties, Computer Science Department, Naval Postgraduate School, Monterey, California, www.laas.fr/IFIPWG/Workshops/44/W1/04_Dinolt.pdf

Heiser J, (July 2002), InfoSecurity Magazine, Security Through ROSI-Colored Glasses, infosecuritymag.techtarget.com/2002/jul/curmudgeons_corner.shtml

Insight Consulting, (undated), CRAMM Version 5 Overview, www.insight.co.uk/datasheets.htm

Kaplan R S, Norton D P, (1996), Harvard Business School Press, Boston, The balanced scorecard: translating strategy into action, ISBN 0-87584-651-3

Knight J C, Sullivan K J, (2000), Towards a Definition of Survivability, Department of Computer Science, University of Virginia, Charlottesville, Virginia, www.cert.org/research/isw/isw2000/papers/27.pdf

Kovacich G L, (1998), Butterworth-Heinemann, Information Systems Security Officer's Guide, ISBN 0-7506-9896-9

McInerney S and Montgomery J (2003), Metrics for Network Enabled Capability, Unpublished MoD report, January 2003

QinetiQ, (2001), Security Requirements Models to Support the Accreditation Process, QinetiQ/KIS/SEB/CP011079/1.1, presented at the Annual Sunningdale Accreditors' Conference 10th-11th September 2001, www.qinetiq.com/home_enterprise_security/conference_papers_index.Par.0004.File.pdf

QinetiQ, (2003), Managing Infosec Risk in Complex Projects, QinetiQ/KIS/TIM/CP010069/2.0, www.qinetiq.com/home_enterprise_security/conference_papers_index.Par.0005.File.pdf

SANS Glossary of Terms Used in Security and Intrusion Detection, Updated May 2003, SANS Institute, www.sans.org/resources/glossary.php

Scholtz T, (2002), META Group, Appropriate Investment in Information Security: Is Risk Assessment Enough? www.metasecuritygroup.com/library/deltas/AppropriateInvestment.pdf

Sharples M, (2002), QinetiQ Ltd., CARE ASAS Activity 2: WP2 System Performance Metrics, www.eurocontrol.int/care/asas/documentation/care-asas-a2-02-033.pdf

Smith Jr. Dr E A (undated), Naval War College Review, Network Centric Warfare: Where's the beef?, http://www.iwar.org.uk/rma/resources/ncw/smith.htm

SSE-CMM Security Metrics, (undated), www.sse-cmm.org/metric/metric.htm

Swanson M, Bartol N, Sabato J, Hash J, Graffo L, (2002), National Institute of Standards and Technology, Gaithersburg, USA, Security Metrics Guide for Information Technology Systems – NIST Special Publication 800-55, csrc.nist.gov/publications/nistpubs/800-55/sp800-55.pdf

Vaughn, Rayford B Jr., (2001), Are Measures and Metrics for Trusted Information Systems Possible? A Position Paper, philby.ucsd.edu/~cse291_IDVA/papers/rating_position/Vaughn.pdf