A New Paradigm for Trusted Systems

Dorothy E. Denning
Computer Science Department
Georgetown University
225 Reiss
Washington, D.C. 20057

1 The Current Paradigm and Breakdown

The current paradigm for trusted computer systems holds that trust is a property of a system. It is a property that can be formally modeled, specified, and verified. It can be "designed into" a system using a rigorous design methodology. For high levels of assurance, the design methodology uses formal models and methods in order to "prove" that trust is present.

This paradigm underlies *The Department of Defense Trusted Computer System Evaluation Criteria* (TCSEC) [3], commonly called the "Orange Book," and its companion "rainbow series" reports. In this paper, we will refer to these documents as the "Criteria." The Criteria specifies a methodology for modeling, designing, and implementing a system that builds trust into a system, and a process for proving to an evaluator that the methodology has been followed. For a description of the Criteria and the evaluation process, see Chokhani [1].

Application of the Criteria has been fraught with problems for both developers and evaluators. Steve Lipner clearly articulated this breakdown in the keynote address at IFIP-SEC 91 [4]. The problems he identified include:

1. Systems are not operated in their evaluated configuration. Evaluated systems are penetrated because they are not properly configured or operated.

2. The Criteria apply to operating systems products, whereas actual operating environments include heterogeneous networks and applications.

3. Applications must run with "privilege," overriding the operating systems controls. Evaluation becomes irrelevant. There is no experiential basis on which to build application-level criteria.

4. Real systems are vastly more complex than their security models. The vendors learn what system settings, tools, and documentation are needed from the experiences of their customers with their products.

5. The security management documents are thick and there is a forest of controls. The paperwork required of vendors is an enormous burden.

6. By the time a product has been evaluated, it is obsolete.

7. The Rating Maintenance Program (RAMP), which was designed to allow vendors to self-evaluate new versions of a product, imposes a plethora of paperwork, checking, bureaucracy, and mistrust on vendors.

8. No one knows what a class C2 system is. Part of the problem lies in applying an abstract model of subjects and objects to real systems when it is not at all obvious what should be subjects and objects in the system.

Because of these problems, it has been necessary to produce "interpretations" of the Criteria. The interpretations grow and change as new systems are evaluated, but nonetheless remain ambiguous.

Lipner offers some suggestions for improving the process. While his suggestions are likely to alleviate some of the problems, we propose that we also rethink the question "What is a trusted system?" My initial investigation into this question suggests that the current paradigm, which treats trust as a property, is inconsistent with the way trust works in the world. By shifting to a paradigm that is consistent with the realities of trust, we may be able to produce trusted systems at considerably reduced cost, effort, and aggravation. We shall propose such a paradigm here, and we invite the reader to explore its implications with me.

The need for a paradigm shift is not limited to the domain of security. Peter Denning [2] has noted that software quality is held as a property that can be designed into a system by a four-stage process: formulate the requirements, develop formal specifications for the requirements, develop programs from the specifications, and demonstrate that the programs meet the specifications. He proposes a shift in paradigms by reframing the question "What is software quality?" to "How do we satisfy the customers of our software?"

The paradigm for trusted systems presented here similarly focuses on producing systems that satisfy customers, in this case, systems that customers trust in the domain of security.

2 What is Trust?

2.1 Trust is an Assessment

The word "trust" is used with people, organizations, and objects. It is an assessment that a person, organization, or object can be counted on to perform according to a given set of standards in some domain of action. As an assessment, it is a declaration made by an observer rather than an inherent property of the person, organization, or object observed.

For example, we may trust a person to speak truthfully, keep promises, arrive on time, give an entertaining talk at a conference, represent our concerns at an important meeting, lead a project, implement a program, fly an airplane, or perform open heart surgery. We may trust an organization to keep our records confidential, deliver certain types of products or services, or refund our money if we are unsatisfied. We may trust an airplane to not crash, a bridge to not collapse, the groceries we purchase to not be contaminated or poisonous, or a program to perform its stated function and not have undesirable side effects.

An assessment of trust is always relative to a domain of action. We may trust a person to give a stimulating lecture on computer crime, but not trust them to fly an airplane or cook a Thai dinner. We may trust a woodworker to produce a cabinet of exceptional quality, but not trust them to deliver it on time. Thus, people are not simply trusted or not trusted, but rather trusted or not trusted in a particular domain. However, we often lose the distinction of domain, generalizing assessments of trust across domains. For this reason, we often hear people say things like "This person cannot be trusted."

Likewise, an assessment of trust is always made against a set of standards in the domain of action. These standards evolve in communities of people who interact and coordinate action together, and they may differ from one culture to the next. They are often loosely defined or subjective, for example, standards for a "good teacher," a "good restaurant," or a "good department." They may be so ingrained in our culture that we are not even consciously aware of their presence. Yet they play a critical role in our coordinated actions in the world.

The domains and standards for trust change over time as new technologies come to market and new breakdowns occur. A few years ago, nobody was concerned about whether a floppy disk might contain a computer virus or other form of malicious code. Now people are reluctant to trust a disk if they are not sure of its origin. The Tylenol scare led to higher standards for packaging drugs and other goods.

An assessment of trust may or may not be grounded. It is grounded if evidence can be produced that the standards are met. Otherwise it is ungrounded. In many situations, it is less important whether an assessment is grounded than whether it is believed. People act out of their beliefs even when there is no evidence to support them.

2.2 How Assessments of Trust are Made

We make assessments of trust based on our experiences in the world. As we interact with other people, organizations, and objects, we observe the effects and form our assessments. If a person consistently keeps their promises, then we trust that person to keep future promises. But if they fail to keep a promise, we may begin to distrust them. Similarly, if we try a new restaurant and have a good experience, then we may make an assessment that the restaurant is excellent. However, if we go back and have a bad experience, we will change our assessment and possibly never return. We often make assessments of trust based on a single incident; this is why first impressions are so important.

If we do not have direct experience with a person, organization, or object, we will make an assessment of trust based on the declarations of others whom we trust. If a person whom we trust says that another person is an entertaining and stimulating speaker, then we may accept their assessment and invite the person to give a talk at a conference. If a restaurant critic or friend reports on a new restaurant, then we may use their assessment to determine whether to try the restaurant. If a popular computing magazine reports that a particular vendor provides better service than a competitor, we may decide to order products

from that vendor. We make purchases, hiring decisions, travel plans, invitations, and other decisions based on what others say when our own experience is inadequate.

There has been a growing industry relating to the buying and selling of assessments of trust. This industry includes organizations such as Consumer Reports; consultants and consulting firms with expertise in specialized domains; and magazines, newsletters, and articles which evaluate products, services, and organizations. Although we often rely on the assessments of others, we give greater weight to our own experiences, and we will not accept another person's assessment if it contradicts our own experience. Instead, we may lose trust in the other person's assessments. We are most influenced when we lack experience of our own.

We thus ground our assessments of trust on our personal experiences and on the experiences of others whom we trust. We seldom base our assessments on mathematical theories. The Golden Gate Bridge is trusted, not because someone proved mathematically that it would not collapse, but rather because it has withstood over 50 years of service. In 1987, it passed an impromptu "proof test" by supporting the largest load ever, 250,000 people. By comparison, the Tocoma Narrows Bridge, which was built using the same theory, was destroyed by wind in 1940 [5,6].

This does not mean that formalism has no role in the establishment of trusted products. Formal theories and methods may be used to validate certain aspects of a product, e.g., to show that a circuit design or software module will satisfy certain properties. These methods can help the developers establish trust in their product before it is released. However, the product itself will be assessed by users according to their standards. If a software product shows no evidence of containing malicious code after several years of use, then it will be trusted to be non-malicious regardless of whether that property was formally proved.

2.3 Trust is a Critical Element of Markets

Assessments of trust are thus formed and shared in a world where we interact with the people, organizations, and objects around us. This world is also a marketplace of transactions, and the value of a person, organization, or object in the market will be determined to a large part by the amount of trust that others have in them. If a person has a reputation of being a highly talented athlete and of high integrity, then that person will have many opportunities in the market. Similarly, if a service provider has a reputation of providing exceptional service at competitive prices, then it is likely to do well. But reputations are volatile. Once a person or organization acquires a reputation of being untrustworthy, it can be hard to overcome that reputation even if the assessment was poorly grounded.

The word "market" is being used in a very loose way to refer to the space of all transactions, including social transactions that do not involve money. A transaction is any exchange between two parties. The transaction may involve loaning a book in exchange for the right to borrow one in the future or even for the friendship that will follow from the loan. A conversation can be regarded as an exchange where two people share information, beliefs, thoughts, and emotions.

In this market, people can trade as they choose, subject only to their own ability to make offers that are desired by others, and by the regulations and rules that are imposed by governments and private organizations. The viability of a person, company, or product in the world is strongly determined by the trust they evoke in those they wish to interact and trade with. The market will eventually weed out people, organizations, and products that are considered untrustworthy, though this may take time if there is little or no competition in that domain. In a sense, the market determines the criteria for trust based on the needs and demands of the people.

In the domain of aircraft, for example, the market has demanded planes that do not crash. If a plane crashes and the cause of the crash can be attributed to a design flaw, then people will not fly on planes of that type. This happened to the DC-10 after one incident, and there are people who still avoid it.

3 The New Paradigm

The current paradigm of treating trust as a property is inconsistent with the way trust is actually established in the world. It is not a property, but rather an assessment that is based on experience and shared through networks of people in the world-wide market. It is a declaration made by an observer rather than a property of the observed.

In the new paradigm, we see that a "trusted system" is one that produces assessments of trust. These assessments are based on standards of performance and are grounded in observable behavior of the product in the marketplace. The standards for trust will change as new technology, new threats, and new practices are introduced in the market. Moreover, the assessments about a particular system will be continu-

ally remade each time the system is used. Ultimately, a system is trusted if and only if its users trust it.

The new paradigm has several implications relating to the Criteria and to producing trusted systems. The following touches briefly on these implications. Further study is needed to develop a more complete understanding of the proposed paradigm shift.

3.1 Security Criteria

At first glance, it might appear that the current Criteria recognizes that trust is an assessment rather than a property since the security rating assigned to a system (C2, B1, etc.) is an assessment. However, the Criteria are based on the assumption that trust is a property that can be built into a system following specified design methodologies rather than the premise that trust itself is an assessment made by users based on how well the observed behavior of the system meets their own standards.

In the new paradigm, security criteria would articulate the (possibly unstated) standards that users employ when making assessments of trust; that is, they would formulate the concept of customer satisfaction in the domain of security. They would emphasize those features that customers are most concerned about, for example, protection against break-ins and viruses, simple access controls, ease of use, and product support.

Since users do not particularly care how a system is structured internally or the methodologies used during development, the security criteria would not specify how a system should be modeled, structured, designed, or developed as in the current Criteria. For example, there would be no concept of security kernel, trusted computing base (TCB), or formal security policy model. There would be no requirements on system architecture, design specification and verification, or configuration management.

The standards would be specific to different types of products and stated in terms of actual users, processes, and entities rather than abstract subjects and objects. Thus, they would not require "interpretation" of an abstract security model and they would be readily understandable to users and developers alike.

To illustrate, the standards for operating systems might include discretionary access at the level of individual files and users, logging of all successful and failed login attempts, and break-in prevention. The standards for database systems might include discretionary access at the level of records, attributes, and individual users, and logging of all database accesses at the relation level and all updates at the record level. The standards for virus protection software might include the ability to detect any virus in a specified list and the ability to remove any detected virus. The standards for networks and communication systems might include optional encryption using the Data Encryption Standard.

There may be a common set of standards applicable to all types of products, for example, standards for product service and support. Since many security problems arise from improper installation or operation, or from flaws that are discovered after the product has been released, product support is a significant factor in customer satisfaction and assessments of trust.

The standards might be classified according to whether they are required for a certain "level of security" or for certain types of environments (banking, hospital systems, etc.). For example, being able to withstand penetration attacks from legitimate users might be associated with a higher level of trust than preventing break-ins. A product could be evaluated by checking off the standards that it meets.

"Security benchmarks" could be included with some of the standards. For example, consider a standard for break-in prevention. This standard could be assessed through a "break-in benchmark" that could be run against a system to see if it succumbs to certain attacks, for example those that use password cracking programs or exploit potential network protocol vulnerabilities. One can envisage other benchmarks, for example, to assess the ability of a virus protection package to detect viruses.

This approach of assessing observable behavior and of using benchmarks is not new. Indeed, it has arisen naturally in the market in response to customer needs. There have been many published articles that rate or compare security packages in concrete terms, and vendors and researchers have developed software tools that can test for the presence of weak passwords, improper defaults and system settings, and various other vulnerabilities. All of these assessments and tools have been developed with the goal of meeting the needs of customers, and are entirely consistent with the way trust works in the world. Thus, the paradigm described in this paper is already practiced in the commercial world, and the existing practices provide a useful starting point for determining security criteria.

Although the Criteria is based on a model of trust that is inconsistent with the way trust works, it offers much towards the construction of new security criteria. Many of the requirements relate to functionality needed by users, and while many are abstract,

they could be made concrete. The requirements for penetration and covert channel testing identify areas where benchmarks could be created, although it is unclear that protecting against most covert channels corresponds to any real-world market need. The security criteria would be driven by market forces. They would reflect the current standards for trust in the market, and they would change with market needs. They would be developed by or at least with users representing a variety of different customer bases.

Although there could be more than one set of standards, a national or international standard has the advantage of providing industry with a clear set of guidelines. The standard(s) could be produced by the government through the current NIST/NSA effort or by other standards groups, for example, ANSI, the IEEE, and ISO.

Although security criteria articulate community standards for trust, a system that meets the criteria is not necessarily trusted. Ultimately, trust is always determined by users whose needs may deviate from the community standards. This underscores the importance of product support from a vendor.

3.2 Producing Trusted Systems

In the new paradigm, vendors would be free to design and develop systems using any architecture and methodology they choose. The security criteria would not impose any particular structure or methodology on the customers. Security kernels, formal models and methods, and other developmental requirements in the current Criteria would be used only to the extent that vendors perceive that the return on their investment justifies the cost. The requirements in the current Criteria, coupled with the costly evaluation process, have led many vendors to conclude that it is simply not worth the effort to develop systems at those levels where formal methods are required. Removing these requirements opens up the possibility of considerable innovation in the development of trusted systems. Researchers may be able to uncover structures and methodologies that produce trusted systems at considerably reduced cost.

The current Criteria were developed with the objective of eliminating all security risks, at least at the higher levels. By adapting a particular architecture and following a specified design methodology based on formal specifications and proofs, security risks would be avoided. This risk-avoidance strategy has the disadvantage of inhibiting innovation and progress in system architecture and development. If followed to its extreme, it will guarantee that "trusted systems" are archaic and not cost-effective. As illustrated by Petroski [6], progress in engineering comes only when designers take risks. Taking risks is essential in order to build systems that are more economical, functional, or aesthetically pleasing than their predecessors. Moreover, we learn more from our failures than our successes, and progress depends on failures. A strategy of creating criteria that eliminate security risks is especially dangerous because we lack worked examples, especially for applications such at database systems, transaction processing sysems, and heterogenous networks. A better strategy is to encourage risk taking while disseminating knowledge about failures through channels such as CERT and security publications.

4 Summary

The current paradigm for trusted systems holds that trust is a property of a system. We have argued that this paradigm, which underlies the Criteria for trusted systems, is inconsistent with the way trust works in the world.

We then examined the concept of trust, showing that trust is an assessment made by an observer about a person, organization, or object observed. These assessments are formed and shared in a world-wide market where people interact with each other, with organizations, and with objects. Our own assessments are based on our personal experiences and on the assessments of others whom we trust.

This understanding of trust as an assessment formed in a market leads to a radically different approach to the development of security criteria. In this paradigm, the criteria would be a set of standards directly related to customer satisfaction. The standards would reflect current market requirements, be specific to different types of products, and be stated in terms of actual users, processes, and entities rather than abstractions such as subjects and objects. They would continually evolve to respond to new technologies, new threats, and new demands in the market.

The criteria would not impose requirements on the internal structure of a system or on development methodologies. The vendors would be free to choose their own methods for producing secure systems. Their systems will be evaluated according to market-based criteria for customer satisfaction, and they will be trusted as long as they meet the evolving standards and needs of the customers.

Further study is needed to determine whether the proposed approach is sound for at least commercial

systems if not military ones. If it is, then additional work is needed to identify the current community standards in order to formulate new criteria. Beyond that, the approach opens up the possibility of new security architectures and methodologies, and of news products that support the evaluation process, in particular security benchmarks.

Acknowledgments

We are grateful to Peter Denning, Hilary Hosmer, Bob Lawton, and Steve Lipner for their generous and helpful comments on an earlier version. Once again, my colleagues saved me from at least one round of public embarrassment.

References

[1] Chokhani, S., "Trusted Products Evaluation," *Comm. of the ACM*, Vol. 35, No. 7, July 1992, pp. 64-76.

[2] Denning, P. J., "What is Software Quality?" *Comm. of the ACM*, Vol. 35, No. 1, Jan. 1992, pp. 13-15.

[3] Department of Defense, "Trusted Computer System Evaluation Criteria," DOD 5200.28-STD, December 1985.

[4] Lipner, Steven B., "Criteria, Evaluation, and the International Environment: Where Have We Been, Where Are We Going?" Proc. IFIP-SEC '91; also in RISKS-FORUM Digest 12.46, October 1991.

[5] Petroski, H., "Making Sure," *American Scientist*, Vol. 80, March-April 1992, pp. 121-124.

[6] Petroski, H., *To Engineer is Human, The Role of Failure in Successful Design*, Vintage Books, 1992.

They devised a system of government that was revolutionary for their time in response to the revolution of their time. Though it was far from perfect and it encountered many growing pains along the way, it proved to be an effective and efficient system that managed to serve a wider spectrum of the population than the old ways of feudalism. That spectrum gradually expanded to include more people, and it eventually broke the chains of slavery as the ideas of the Enlightenment finally reached a critical mass a century later (and yet a larger one a century after that, i.e., the Civil Rights Moveme The current paradigm for trusted computer syst,ems holds that trust is a property of a system. It is a. property that ca.n be formally modeled, specified, and verified. It can be â€œdesigned int,o â€ a. system using a rigorous design methodology. For high levels of a.ssurante, the design methodology uses forma.1 models sod methods in order to â€œprove â€ tha.t trust is present.Â The Alliance intends to develop â€œa new computing platform for the next century that will provide for improved trust in the PC platform.â€ They want you to trust these platforms enough to use them for E-commerce. By promoting the concept of a trusted subsystem and chains of trust between those systems, such a platform has a good chance of becoming the basic building block for electronic commerce. View. A New Paradigm for Trusted Systems. Dorothy E. Denning Computer Science Department. Georgetown University 225 Reiss.Â The current paradigm for trusted computer syst,ems holds that trust is a property of a system. It is a. property that ca.n be formally modeled, specified, and verified. It can be â€œdesigned int,oâ€ a. system using a rigorous design methodology. For high levels of a.ssur-ante, the design methodology uses forma.1 models sod methods in order to â€œproveâ€ tha.t trust is present. This paradigm underlies The Department of De-fense Trusted Computer System E~valuation Criteria (TCSEC) [3], commonly ca.lled the â€œOrange Book,â€ and its companion â€œrainbow seriesâ€ rep0rt.s. In this paper, we will refer to thes