IDA

INSTITUTE FOR DEFENSE ANALYSES

On The Limits of Test in Assuring the Integrity of Products

Brian S. Cohen, *Project Leader*

Kathy Lee

1 April 2014
IDA Non-Standard
NS D-5086
Log: H 13-001819
Copy

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About This Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract DASW01-04-C-0003, Task DD-5-2635, "Trusted Microelectronics," for Defense Microelectronics Activity (DMEA). The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

On The Limits of Test in Establishing Products Assurance

Brian. S. Cohen and Kathy Lee
Information Technology and Systems Division
The Institute for Defense Analyses
4850 Mark Center Drive, Alexandria, Virginia, USA 22311
Contact author email: bcohen@ida.org

Abstract: *Testing is being employed by DOD as one defense against selected exploitations of supply chains, with policy and practice calling for testing to detect counterfeit and tampering of parts. The limits of testing for reducing these particular risks is explored, and the results show that testing works best for simple low quality parts, but poorly for complex high quality parts. This suggests that testing will be less effective as a primary means of managing the risks of counterfeit introduction and tampering with parts when compared to other means such as using trustworthy suppliers (such as a Trusted Supplier accredited by DMEA).*

Keywords: counterfeit; acceptance testing; risk management; assurance; inspection.

Introduction

Significant emphasis is being placed on incoming acceptance testing as a practice for detecting counterfeiting and exploitation in the supply chains for defense systems. Testing has been identified as one of the primary mitigations in recent defense policy, with the Trusted Systems and Networks policy [1] requiring programs to "detect vulnerabilities within custom and commodity hardware and software through rigorous test and evaluation capabilities, including developmental, acceptance, and operational testing" in critical components. Further, the new counterfeit prevention policy [2] calls for the defense enterprise to "detect counterfeit materiel using sampling techniques, materiel testing, and auditing." While significant resources are being directed to, and dependence is being placed on, testing as a defense against these supply chain exploitations, this paper explores the limits of testing as a means of detecting counterfeiting and tampering. The discussion will use counterfeiting as a way of understanding the problem, but the results could also apply to tampering. The end result of this analysis is the conclusion that testing can be a cost-effective means of managing risk for products either of low quality or having high rates of counterfeiting/tampering, but for products whose anticipated counterfeiting/ tampering rates are very low already, acceptance testing alone may be an extremely counterproductive means of improving the detection of counterfeiting or selected forms of tampering.

Two important dimensions of the problem are considered. The first examines the effectiveness of testing (in managing risk) in the screening of "lots," and the second examines the effectiveness of screening within a "lot." The first dimension is critical when evaluating whether the potential increased cost of purchasing from a trustworthy source (such as an original manufacturer or a Trusted Supplier) is better than purchasing from an untrustworthy source and using testing to establish product assurance. The second dimension considers purchased lots that may actually have been tainted by "salting," in which some individual parts are counterfeit or have been tampered although the majority of the lot comprises authentic pristine parts. In the remainder of this paper we will discuss counterfeits, but the entire discussion applies to both parts that are counterfeit and those that have been tampered.

This paper examines the effectiveness of testing techniques when applied as a screening process during the purchase process for components. Figure 1 provides a flow chart for screening for product assurance.

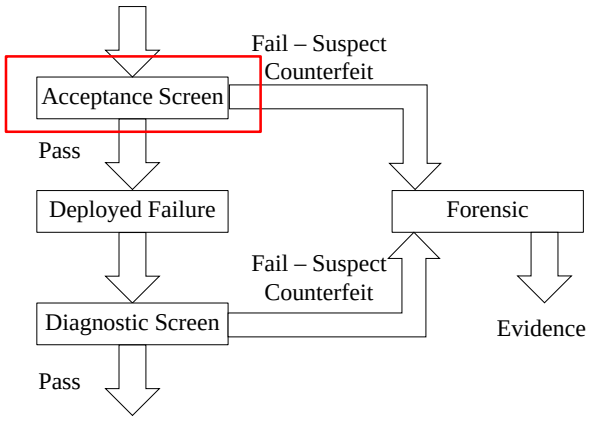

Figure 1. Using Test as a Screen for Product Assurance

A screening process will typically classify products as being "good" or "bad." In this context, we are applying the screening process to classify a product as either counterfeit or authentic. A product that is found non-conforming is considered a counterfeit. We use the term "suspect counterfeit" to differentiate the result of the screening from the actual ground truth or the conclusions of a legal finding. Figure 2 captures the classification problem for screened counterfeit and original parts.

		Actual	
		Counterfeit (C)	*Original (O)*
Screening	*Suspect Counterfeit (SC)*	True Positive (TP)	False Positive (FP)
	Suspect Original (SO)	False Negative (FN)	True Negative (TN)

Figure 2. The Classification Problem

Evaluating Risk Mitigation for Lot Purchases

Risk is essentially the expected value accounting for the cost of acquiring a component and the costs associated with negative impacts from exploitation (there may even be additional positive benefits that will need to be accounted for). For the purposes of this analysis, we don't account for the actual purchase price of the component; we focus on the added costs of testing. While it is true that purchase prices for components from trustworthy and untrustworthy sources differ, we don't focus on that issue and do argue that when a component is available from a trustworthy source the price markup is usually trivial. In the following discussion, we use the following definitions of various costs:

Counterfeit Cost (CC) – Cost of an actual counterfeit lot being used and causing an adverse consequence (System fails, requires additional repair, etc.)

Suspect Cost (SC) – Cost of a suspect counterfeit lot which requires mitigation and disposition

Test Cost (TC) – Cost to perform the screening tests

Counterfeit Rate (CR) – Percentage of lots that are counterfeit

In a starting situation, our risk is simply the expected cost of counterfeit escapement consequences:

Risk = CC * CR.

When we apply a screening process, we have the cost of the screening and hopefully a reduced counterfeit escapement consequence. But we also have to account for any costs associated with handing identified suspect counterfeits:

Residual Risk = TC (Screening Cost)

+ CC * FN * CR (Counterfeits Escaping)

+ SC * (TP *CR + FP * (1-CR))

(Dealing with Suspect Counterfeit).

The cost of a counterfeit "escaping" into a deployed system can be significant – incurring costs to replace bad components and exposing the system to the potential of serious failures. In one example [3], a single counterfeit escapement resulted in $2.7 million (M) in costs to mitigate. This does not include the additional risk that the counterfeit could cause system failure. Some of the other costs associated with risk include the cost of dealing with detected suspect counterfeits which may result in additional forensic testing to gather evidence, validation by the original manufacturer, reporting to the Government–Industry Data Exchange Program (GIDEP), Electronic Resellers Association International (ERAI), and law enforcement, and then appropriately quarantining, marking, scrapping, or returning the suspect items. Each suspect counterfeit might result in as much as $1 thousand (K) in added costs, a not insignificant amount. Finally, when balancing the costs and benefits, it is important to consider the actual cost of screening, which based on rough order estimates from defense primes and test houses can range between $2K for a simple AS5553 [4] incoming test to $7.5K for an AS6081 [5] outsourced testing.

Another significant factor that affects risk is the underlying rate of counterfeiting. In most cases, we assume that all the parts in a purchased lot are from the same underlying population. In this way, the counterfeit rate is the percentage of lots which contain one or more counterfeits. We will later examine the problem of determining whether a lot actually comprises the same population or suffers from "salting." The Defense Logistics Agency (DLA), which buys mostly obsolete components, encounters some of the highest risks from counterfeits and through its testing program found counterfeit rates that were at least 1.3%, and by some measure might be as high as 8.6%, in its purchases [6]. On the other hand, purchases made from original equipment manufacturers (OEMs) or through authorized distribution likely have counterfeit rates of much less than 0.01%. Of greatest concern is that the vast majority of the components entering the defense supply chain lies between these two ranges; this middle range will be the focus of later examination.

Current Screening Performance

A recent round robin test [7] provided 11 test labs with 2 parts, one counterfeit and one authentic, and had each lab evaluate each part's authenticity. The result was far from perfect, with one lab incorrectly identifying the counterfeit part as authentic and two labs incorrectly identifying the authentic part as counterfeit. Another lab was unable to state a conclusion. The resulting matrix of detection rates is show below in Figure 3.

		Actual	
		Counterfeit (C)	***Original (O)***
Screening	***Suspect Counterfeit (SC)***	**TP: 82%**	**FP: 28%**
	Suspect Original (SO)	**FN: 18%**	**TN: 72%**

Figure 3. Example of Screening Performance

The Problem of False Positives

False positives occur when an actual authentic component is incorrectly identified as a suspect counterfeit. This can happen in practice because tests simply have variable

results, but it can also happen because the reference or golden sample is poorly characterized. And in practice, the component buyer has far less knowledge about the product than the original manufacturer, they don't have test vectors, and they have very little ability to determine when differences and anomalies should raise suspicions about counterfeiting or tampering.

In the round robin example above this rate was very high at 28%. If we plug these example numbers into the risk formulation presented earlier we find that for the case where we are buying obsolete components and have 1.3% of the products being counterfeit (the lower end of the DLA observed rate), we can reduce the risk from:

Risk = $2.7M * 1.3% = $35,100

to:

Residual Risk = $2K + $2.7M * 18% * 1.3%

+ $1K * (82% * 1.3% + 28%* (1-1.3%))

= $8,605.

Clearly, in this case the application of screening can significantly reduce the risk. But when we have a much lower counterfeit rate, such as 0.01% we find that the starting risk is lower:

Risk = $2.7M * 0.01 = $270

And the Residual risk is:

Residual Risk = $2K + $2.7M * 18% * 0.01%

+ $1K * (82% * 0.01% + 28%* (1-0.01%))

= $2,328.

It is clear that for this case, the risk is already modest and that screening is actually counterproductive. There are two key problems – first, the screening cost itself is far more expensive than the expected value of a counterfeit escaping in the first place, and second, when we compute how much it costs to deal with suspect counterfeits, more money is spent dealing with the false positives ($280) than the expected risk cost.

Evaluating Risk Mitigation within a Lot

We mentioned earlier that parts within a lot may actually be "salted," with bad parts inserted into a lot of otherwise good components. This scenario, which has been observed specifically involving counterfeits, is a particular problem in that it undermines a fundamental assumption in traditional quality practices. Traditional quality practices have assumed that all parts in a lot with a single date/lot code are from the same underlying population of parts and thus of the same quality.

When we cannot make this assumption, we may have to test a large number of components in order to achieve a desired assurance level. In the following discussion, we will relate assurance level to quality expressed in parts per million (ppm). Modern integrated circuit manufacturers are typically able to deliver components with quality levels in the 100 ppm range. Manufacturers no longer attempt to use testing to establish quality and reliability, and there is a basic understanding that you cannot test in quality. Instead, buyers employ screening to verify the quality claims from the manufacturer. The challenge in using screening to test in quality is highlighted in Figure 4, which shows that in order to establish quality levels in the 100 ppm range, more than 23,000 components need to be tested. Given that lot sizes are typically quite small (analysis of DLA purchases found that the typical lot size was 168 components [6]), it is likely that even with 100% testing it will be possible to achieve only something on the order of 10,000 ppm quality, far short of the 100 pm we started with. That being said, other data suggests that in obsolete segments of the market where DLA and others are forced to purchase parts from the aftermarket, the baseline quality is much less than 100 ppm. Information from Integra [9] on some 20,000 lots of OEM parts over a decade suggests that in the obsolete aftermarket quality levels range from 1,000 ppm to 25,000.

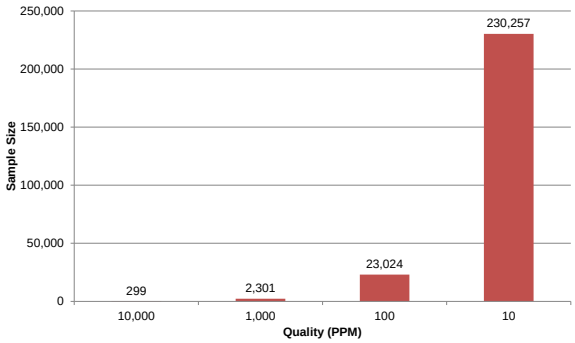

Figure 4. Sample Size needed to establish 90% Confidence in Quality [8]

And a real problem caused by low quality is its effect on reliability. Many long-term reliability factors cause drift in performance as the components age, and when the components have significant variability initially, they may fail prematurely even though 100% of them passed initial testing. Figure 5 shows the relationship between initial quality (yield) and reliability. A 10,000 ppm quality level corresponds to a yield of 0.9, and clearly there are situations in which this has a significant impact on reliability.

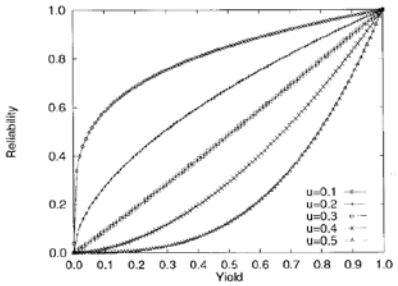

Figure 5. The relationship between reliability and yield of Poisson model for *u* values [10]

Findings

Testing can improve the assurance of lots, but only for sources of supply with high rates of counterfeiting.

Testing lots from sources with low rates of counterfeiting is unlikely to be effective. In many contexts of use, test costs are usually so expensive that they cannot be justified based on selected risk reduction factors. And most importantly, the percentage of authentic parts identified as suspect counterfeit is far too high, causing inordinate handling expenses for these falsely identified suspect counterfeits.

Testing is unlikely to be an effective substitute for using a trustworthy supplier

Testing for acceptance screening is unlikely to be a cost-effective means of assuring that product purchased from an untrustworthy supplier is of comparable assurance to product obtained from a trustworthy supplier. Testing is costly and requires large lot sizes to approach the 100–500 ppm levels of quality of product from the original manufacturer.

Future research questions

The effect of false positives upon the effectiveness of counterfeit test and screening methods is still not well understood. The impact of non-trivial false positive rates may really affect the process when counterfeits are rare and where mitigation and disposition of suspect counterfeits is costly.

When there are quality issues in the supply chain, it is not clear what the role of the quality practices are and how to differentiate quality problems from counterfeiting problems since they are often confounded.

Even when testing cannot assure even modest levels of quality, it is not clear that it can ever assure freedom from counterfeits, let alone tampering. With small lots, 100% testing cannot establish even 10,000 ppm quality, which leaves systems at risk from low reliability issues that can cause premature failure.

And of most importance, a more refined understanding of when to apply testing as opposed to using trustworthy supplies is needed; in particular, a way of taking into account the particular costs and rates which affect the decision needs to be understood and accounted for in the decision model.

Acknowledgements

The authors wish to acknowledge Assistant Secretary of Defense for Research and Engineering (ASD(R&E)), Defense Microelectronics Activity in particular for their support of this work. Additional support for much of the background for this work was provided by ASD(R&E), Systems Engineering, and the Department of Defense – Chief Information Officer. And the authors wish to thank the members of the Society of Automotive Engineers G19 for many stimulating discussions and assistance in the evolution of these ideas.

References

1. DODI 5200.44, “Protection of Mission Critical Functions to Achieve Trusted Systems and Networks,” Issued 11/5/2012 http://www.dtic.mil/whs/directives/corres/pdf/520044p.pdf.
2. DODI 4140.67, “DoD Counterfeit Prevention Policy,” April 26, 2013 http://www.dtic.mil/whs/directives/corres/pdf/414067p.pdf.
3. SASC Report, “Inquiry into Counterfeit Electronic Parts in the Department of Defense Supply Chain,” March 21, 2012 http://www.gpo.gov/fdsys/pkg/CRPT-112srpt167/pdf/CRPT-112srpt167.pdf.
4. SAE AS5553, “Fraudulent/Counterfeit Electronic Parts; Avoidance, Detection, Mitigation, and Disposition,” January 21, 2013 http://standards.sae.org/as5553a/
5. SAE AS6081, “Fraudulent/Counterfeit Electronic Parts: Avoidance, Detection, Mitigation, and Disposition – Distributors:, Counterfeit Electronic Parts; Avoidance Protocol, Distributors,” November 7, 2012 http://standards.sae.org/as6081/
6. Cohen, B and Lee, K, “Exploring the Limits of Testing for Counterfeit Detection,” *DMSMS 2013*, December 4, 2013.
7. Walters, S., “Status of Counterfeit Detection, Steve Walters,” *ARO/CHASE Workshop on Counterfeit Electronics*, Jan 28-29, 2013
8. PPM Calculator from Statistical Solutions http://www.statisticalsolutions.net/ppm_calc.php
9. Sultan L., “Counterfeit IC Avoidance Techniques for Today’s High Performance Memories and Microprocessors,” *SAE 2012 Counterfeit Parts Avoidance Symposium*, November 2, 2012
10. Kim, T. and Kuo, W., “Modeling Manufacturing Yield and Reliability,” *IEEE Transactions On Semiconductor Manufacturing*, Vol. 12, No. 4, November 1999

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YY)	2. REPORT TYPE	3. DATES COVERED (From – To)
01-04-14	Conference paper	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
On The Limits of Test in Assuring the Integrity of Products	DASW01-04-C-0003
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBERS

6. AUTHOR(S)	5d. PROJECT NUMBER
Brian S. Cohen, Kathy Lee	
	5e. TASK NUMBER DD-5-2635
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882	NS D-5086 H 13-001819

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR'S / MONITOR'S ACRONYM
David Pentrack, Defense Microelectronics Activity 4234 54th Street, McClellan, CA 95652-2100	DMEA
	11. SPONSOR'S / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution Statement A. Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Project Leader: Brian S. Cohen

14. ABSTRACT

Testing is being employed by DOD as one defense against selected supply chain exploitations with policy and practice calling for testing to detect counterfeit and tampering of parts. The limits of testing for reducing these risks is explored, and the results show that testing works best for simple low quality parts, but poorly for complex high quality parts. This suggests that testing will be less effective as a primary means of managing the risks of counterfeit introduction or tampering with parts when compared to other means such as using trusted suppliers.

15. SUBJECT TERMS

counterfeit; acceptance testing; risk management; assurance; integrity; inspection

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			David Pentrack
Unclassified	Unclassified	Unclassified	Unlimited	4	19b. TELEPHONE NUMBER (Include Area Code) (916) 231-1576