RSACONFERENCE2014

FEBRUARY 24 – 28 | MOSCONE CENTER | SAN FRANCISCO

Share.
Learn.
Secure.

Capitalizing on
Collective Intelligence

A Penetration Testing Maturity and Scoring Model

SESSION ID: TECH-W02

Dave Shackleford

Founder & Principal Consultant
Voodoo Security
@daveshackleford

Introduction

- Most organizations conduct some variety of penetration tests today
- Reasons vary:
  - Find holes before attackers do!
  - Prove that security issues exist to skeptical management and raise overall security awareness
  - Verify secure system configurations and/or test new technology
  - Discover gaps in compliance posture and satisfy legal and/or governmental requirements
- The range of different pen test scenarios is staggering, though…
  - And we don't really evaluate how "robust" or "mature" a test is

Pen testing overview…and what's missing

Pen Testing…A Short History

- 1960s: Computer “penetration” first discussed by leading experts, with mention of deliberate tests by professionals
- 1970s: The first “tiger teams” were formed
  - 1971, USAF contracted James Anderson to test time-sharing systems
- 1980s-1990s: Much more focus on securing systems overall
  - Farmer and Venema, 1993: “Improving the Security of Your Site by Breaking Into it”
  - 1999: “Hacking Exposed” and so on.

What’s a Pen Test Look Like Today?

Classic Pen Test Cycle

- Reconnaissance
- Scanning and Enumeration
- Exploitation/Penetration
- Repeat steps 2 and 3
- Reporting

What Most People Do

- Scanning with automated tools
- MAYBE exploit with Metasploit or poke around a bit with some other tools
- Produce a generic report

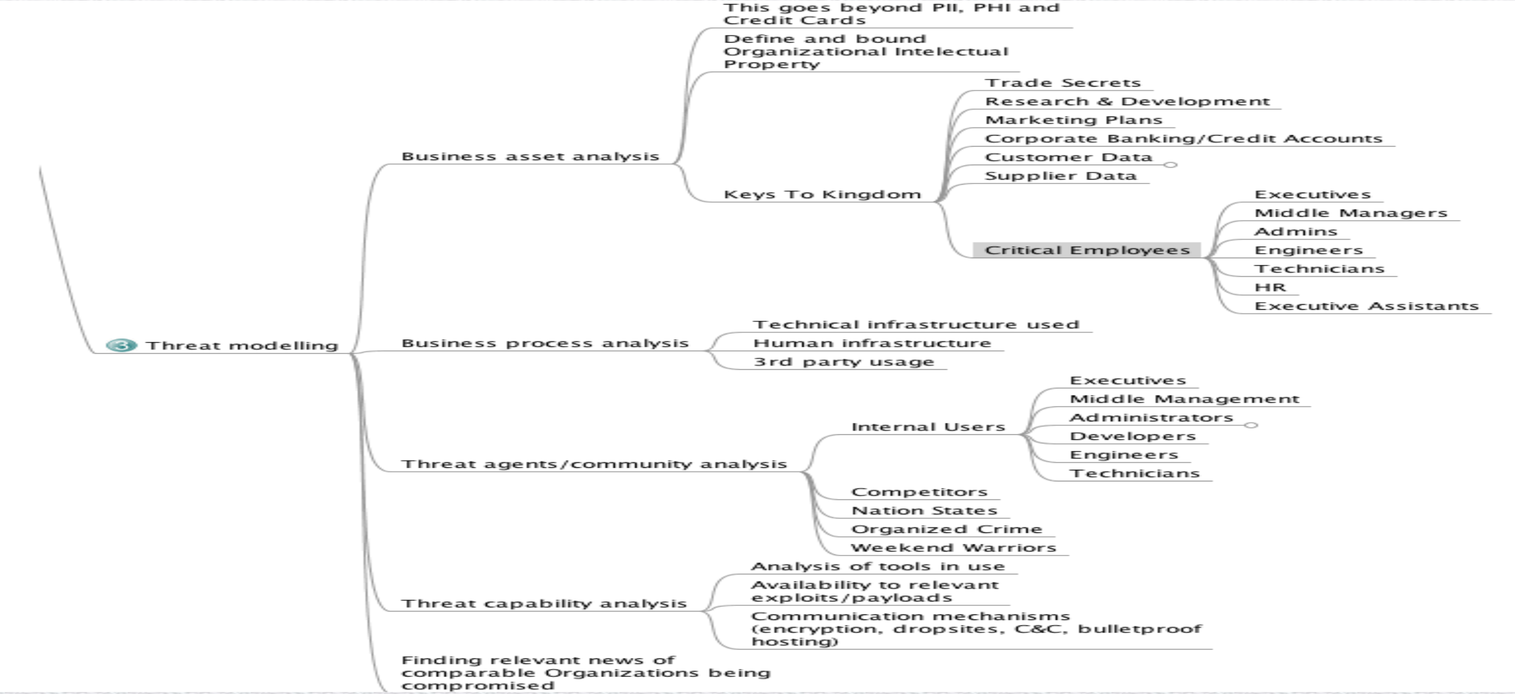
The Problem? We can, and should, do more.

- Skipping reconnaissance is not a good idea. Attackers don’t.
- Running automated tools ONLY will invariably miss things and also produce false positives.
- Many don’t actually exploit…which doesn’t really prove the point
- There’s a lack of realism without stealth and evasion
- Where’s the “people-focused testing” in there?
- And on and on.

Pen Testing Methodologies

- OSSTMM
  - Now v3, more scientific assessment approach
  - www.osstmm.org
- OWASP Testing Guides
  - https://www.owasp.org/index.php/Category:OWASP_Testing_Project
- Penetration Testing Execution Standard (PTES)
  - Includes both detailed instructions and framework diagrams
  - www.pentest-standard.org

PTES Example: Threat Modeling



- 3 Threat modelling
  - Business asset analysis
    - This goes beyond PII, PHI and Credit Cards
    - Define and bound Organizational Intelectual Property
    - Keys To Kingdom
      - Trade Secrets
      - Research & Development
      - Marketing Plans
      - Corporate Banking/Credit Accounts
      - Customer Data
      - Supplier Data
      - Critical Employees
        - Executives
        - Middle Managers
        - Admins
        - Engineers
        - Technicians
        - HR
        - Executive Assistants
  - Business process analysis
    - Technical infrastructure used
    - Human infrastructure
    - 3rd party usage
  - Threat agents/community analysis
    - Internal Users
      - Executives
      - Middle Management
      - Administrators
      - Developers
      - Engineers
      - Technicians
    - Competitors
    - Nation States
    - Organized Crime
    - Weekend Warriors
  - Threat capability analysis
    - Analysis of tools in use
    - Availability to relevant exploits/payloads
    - Communication mechanisms (encryption, dropsites, C&C, bulletproof hosting)
  - Finding relevant news of comparable Organizations being compromised

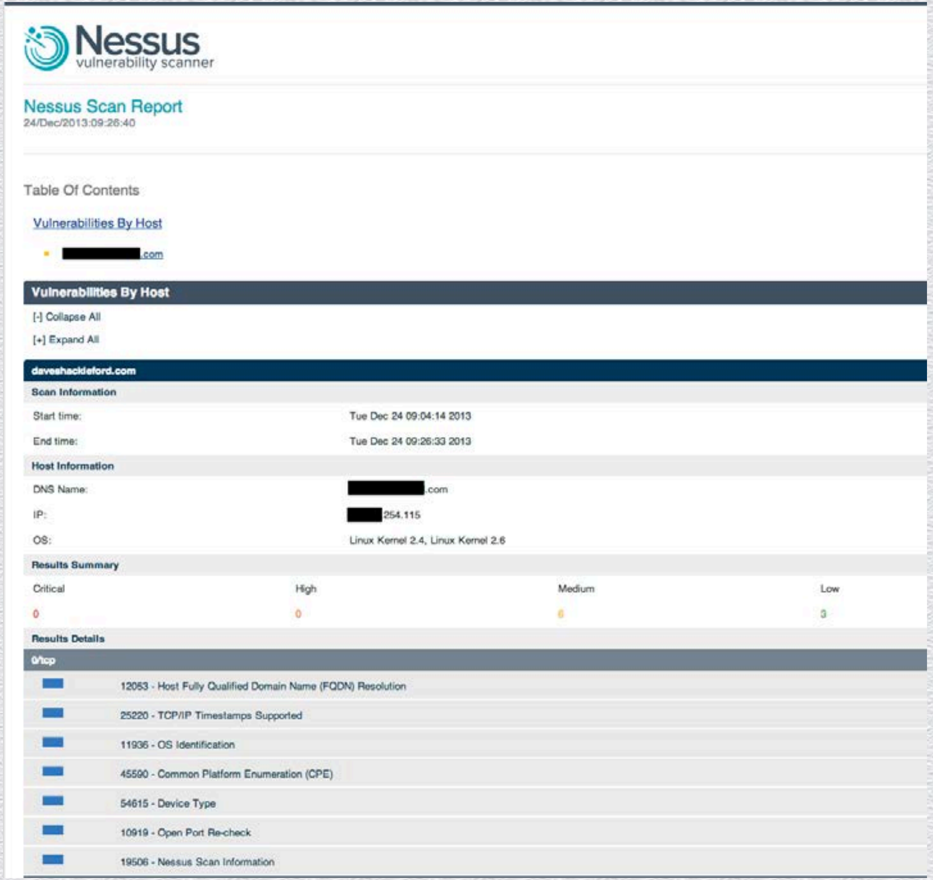
Pen Tests … Minus the Value

- Many scans and pen tests do not really provide as much value as they could
- There are a number of reasons for this:
  - Apathy on the part of your organization, IT Ops, etc.
  - Going for the “checkbox” mentality
  - Unskilled scanning/testing teams
  - Assessment reports that don’t provide tangible benefits
- Lots of ways for a pen test to be “bad”. How can we improve the value and rate a pen test’s maturity?

Common VA/Pen Test Issues

- Copying scan tool data directly into reports, no customization of results
  - Reduces confidence in testing and assessment team
- Lack of concrete remediation guidance
  - This is not really helpful!
- Lack of "repeatability"
  - How was the test performed?
- Inclusion of false positives
  - These should be vetted and removed, or at the least clearly labeled

This is a bad pen test.



Nessus Scan Report

24/Dec/2013:09:26:40

Table Of Contents

Vulnerabilities By Host

- [redacted].com

Vulnerabilities By Host

[-] Collapse All

[+] Expand All

daveshackleford.com

Scan Information	
Start time:	Tue Dec 24 09:04:14 2013
End time:	Tue Dec 24 09:26:33 2013

Host Information	
DNS Name:	[redacted].com
IP:	[redacted]254.115
OS:	Linux Kernel 2.4, Linux Kernel 2.6

Results Summary

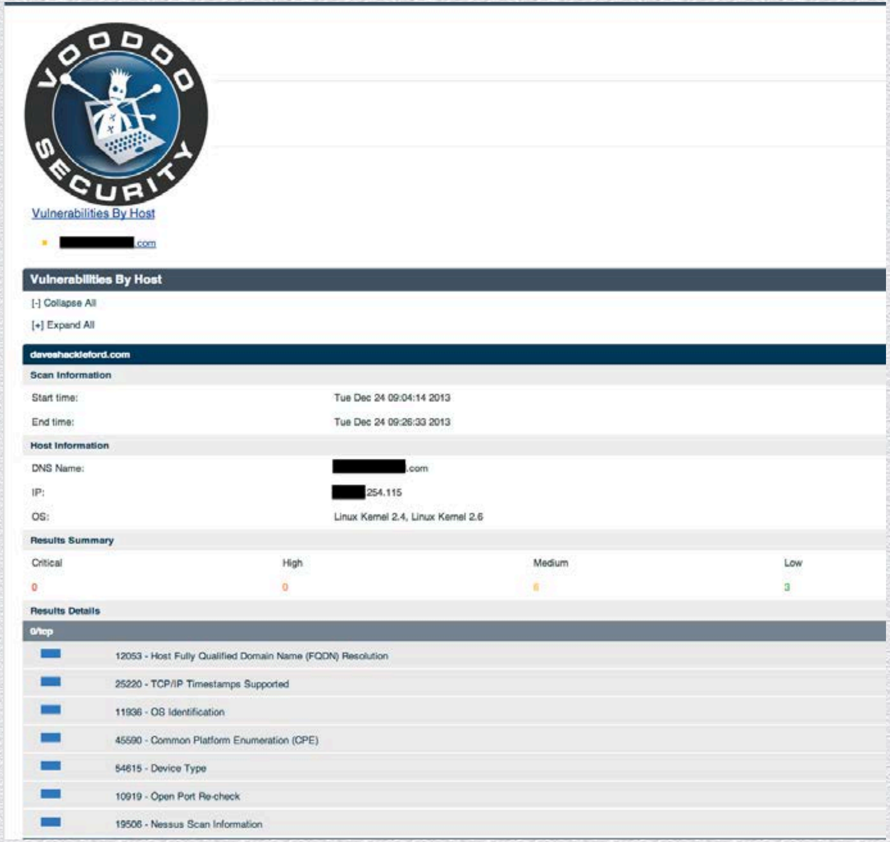
Critical	High	Medium	Low
0	0	6	3

Results Details

0/tcp

- 12053 - Host Fully Qualified Domain Name (FQDN) Resolution
- 25220 - TCP/IP Timestamps Supported
- 11936 - OS Identification
- 45590 - Common Platform Enumeration (CPE)
- 54615 - Device Type
- 10919 - Open Port Re-check
- 19506 - Nessus Scan Information

This is a bad pen test (2)



Vulnerabilities By Host

- [redacted].com

Vulnerabilities By Host

[-] Collapse All

[+] Expand All

daveshackleford.com

Scan Information

Start time:	Tue Dec 24 09:04:14 2013
End time:	Tue Dec 24 09:26:33 2013

Host Information

DNS Name:	[redacted].com
IP:	[redacted]254.115
OS:	Linux Kernel 2.4, Linux Kernel 2.6

Results Summary

Critical	High	Medium	Low
0	0	6	3

Results Details

0/tcp

- 12053 - Host Fully Qualified Domain Name (FQDN) Resolution
- 25220 - TCP/IP Timestamps Supported
- 11936 - OS Identification
- 45590 - Common Platform Enumeration (CPE)
- 54615 - Device Type
- 10919 - Open Port Re-check
- 19506 - Nessus Scan Information

Types of Tests

- “Comprehensive” approach
  - External-based network and/or application tests
  - Internally-oriented network and/or application tests
- The “point-and-shoot” test
  - Testing a specific function/component only, very limited
- The “Standard Operating Procedure” test
  - Part of SDLC or weekly/monthly by internal teams

More Specialized Tests

- Social Engineering
  - If your organization doesn’t allow social engineering, you have a serious blind spot in your threat scenarios
  - Methods include phishing, media drops, tailgating, pretexting
- Red Team tests
  - Can be “anything goes”, ranging from physical tests to social engineering to applications
  - Usually involve data extraction and exfiltration
  - May take much longer, and cost more overall

Maturing Pen Tests & Increasing Value

Getting more value from an assessment

- There are many ways to improve the overall value of an assessment, either internally or by an external provider
- The following are key considerations that should be discussed upfront and understood before an assessment is started:
  - Scoping and rules of engagement
  - Expectations of methodology
  - Results, formatting, and tool output
  - Evidence of compromise and attack vectors
  - Repeatability
  - Prioritized and risk-focused remediation guidance

Scoping and Rules of Engagement

- Establish what networks and apps are in scope
- Determine reason(s) for assessment
  - Compliance
  - General vulnerability understanding
- Rules of engagement include time of tests, how to report vulnerabilities, keeping track of issues, etc.

Have Goals and Targets

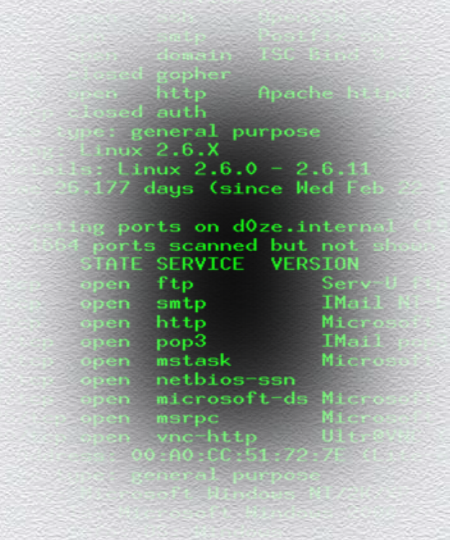
- For pen tests, have a goal:
  - Get to PII
  - Establish specific attack vectors
  - Compromise specific systems or apps
  - Bypass security / stealth attacks
- What is your most sensitive data? What data/access would actually have a material impact?

Expectation of Methodology

- There are not many standards in use today for assessments and pen tests
  - PTES, OSSTMM, etc. mentioned earlier
- All pen testers should have **some** sort of methodology, however.
- This can be as individualized and customized as needed/desired.

Results, Formatting, & Tool Output

- Ensure the assessment report includes details specific to the risks you face and business you are in
  - Don’t be “generic”!
- Demand that external (or internal) testers interpret language from scanners and don’t just paste Nessus results into the report
- Include tools used and output from tools for technical teams to leverage in validation and remediation efforts
  - Usually best to include this as an Appendix or separate report



```
           smtp     Postfix
     open  domain   ISC Bind
   closed  gopher
     open  http     Apache httpd
   closed  auth
  type: general purpose
    Linux 2.6.X
ails: Linux 2.6.0 - 2.6.11
  26.177 days (since Wed Feb

esting ports on d0ze.internal
 1664 ports scanned but not shown
     STATE SERVICE   VERSION
     open  ftp            Serv-U
     open  smtp           IMail
     open  http           Microsoft
     open  pop3           IMail
     open  mstask         Microsoft
     open  netbios-ssn
     open  microsoft-ds   Microsoft
     open  msrpc          Microsoft
     open  vnc-http       UltrVNC
       00:A0:CC:51:72:7E
       general purpose
     Microsoft Windows NT/2K
```

Evidence of compromise & attack vectors

- Demonstrate true compromise or vulnerabilities
  - Screen shots, planted “flags” work best
- Ensure false positives are eliminated
  - Have testers confirm with IT teams before report
  - A “prelim” report may be a good idea for this reason

Realism: Emulating “malware”

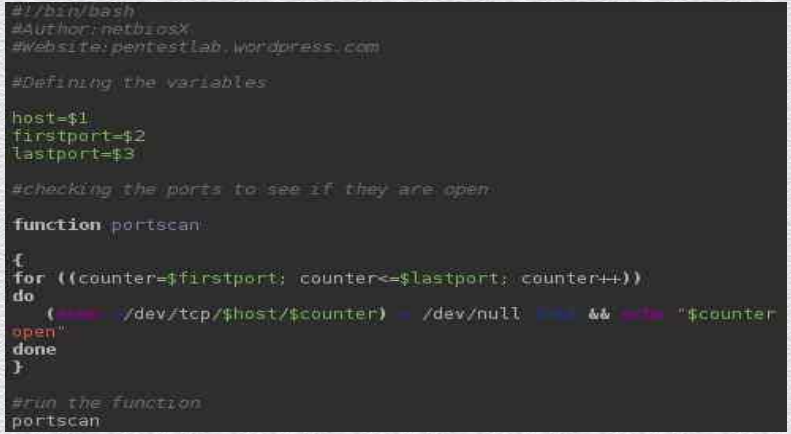
- According to 2013 Verizon DBIR, 40% of compromises involved malware in the year prior
- Grab some malware from one of the following (or elsewhere):
  - http://www.kernelmode.info/forum/viewforum.php?f=16
  - http://contagiodump.blogspot.jp/
  - http://www.malwareblacklist.com/showMDL.php
- Then modify it with an XOR or other encoder
- OR: Write “non-malicious” programs to emulate malware

Malware: Goals

- After exploitation, drop pre-configured malware that **you control**
- This allows for emulation of true attacker behavior
- Ideally, your malware would then connect back to a central control point
  - Depending on the test, you could then pivot to other systems or simply make note of the problem
  - To be more accurate, time infected w/o detection is key

Realism: “Native OS” Attack Vectors

- Fewer attackers are dropping code these days
- Built-in tools on Windows and Linux can be very useful in attacks
- Examples
  - Powershell and WMIC
  - Shell scripting
  - /dev/tcp for socket connections



```
#!/bin/bash
#Author:netbiosX
#Website:pentestlab.wordpress.com

#Defining the variables

host=$1
firstport=$2
lastport=$3

#checking the ports to see if they are open

function portscan

{
for ((counter=$firstport; counter<=$lastport; counter++))
do
   (echo > /dev/tcp/$host/$counter) > /dev/null 2>&1 && echo "$counter
open"
done
}

#run the function
portscan
```

http://pentestlab.wordpress.com/2012/11/12/creating-a-tcp-port-scanner-in-bash/

Realism: Exfiltration

- Exfiltration is covert transfer of information out of a computer system
- Usually accomplished by copying the data from the system via a network channel, although portable media, physical theft, etc. can be effective
- Outbound Channels: HTTP/HTTPS, FTP/FTPS, SSH/SFTP, DNS, ICMP
  - IRC (much less today)
- Continuous/Sporadic

Repeatability

- Listing the actual tools used, the process followed, and how things flowed during the assessment is key
- This can follow the Recon→Scanning→Exploit cycle or some other format
- Good testing firms will tell you what tools they used
  - Include this in the contract
  - Some custom tools may not be handed over at test completion
- This can also be used to validate logs, IDS events, etc.
- Keep track of process with a Wiki or some other tracking tools during assessments

Prioritized and risk-focused remediation guidance

- Define what is important to you in terms of risk
  - Confidentiality for PII and other data?
  - Availability concerns with systems and apps?
  - Integrity with MitM and other attacks?
- Build on this for the report
  - Ensure both attacks and successful exploits are framed in the context of priorities to your business
- Any VA/PT should be focused on your actual risks – not just a scan or exploit to prove you're vulnerable

Risk Rating Descriptions: Basics

Rating	Description
CRITICAL	For CRITICAL findings, the vulnerability should be resolved as soon as possible. Vulnerabilities of this nature expose systems and applications to immediate threat of compromise. Examples include default credentials on Internet-exposed systems or applications, missing critical patches that resolve remotely-exploitable vulnerabilities, and SQL injection attacks that provide access to sensitive data.
HIGH	For HIGH findings, the vulnerability should be resolved within 30 days at most, or as soon as possible. Although these vulnerabilities may entail greater effort for attackers to exploit, they are still very dangerous and may result in successful penetration attempts within a relatively short time. Examples include weak user credentials, privilege escalation attacks, and MitM and other passive attacks.
MEDIUM	MEDIUM vulnerabilities should be resolved within 60 days or as soon as possible. These security flaws may not lead to significant compromise, but could be leveraged by attackers to attack other systems or application components for further damage. Examples include use of plaintext communications protocols, application configuration exposure, and spoofing attacks with minimal data exposure.
LOW	LOW vulnerabilities are largely concerned with improper disclosure of information, and should be resolved within 90 days or as soon as possible. These flaws may provide attackers with important information that could lead to additional attack vectors. Examples include DNS zone transfers, default banners, etc.
INFORMATIONAL	When something is declared INFORMATIONAL it means there is little to no credible threat, but may warrant attention. Examples may include information disclosure on Web sites or applications that is public knowledge.

Risk Management: Context!

- Including relevant information about the target environment that impacts the risk ratings of noted vulnerabilities
  - Commonly include difficulty of exploit, additional controls in place or missing, specific testing circumstances, etc.
- There are two options for leveraging contextual data:
  - Information you have gleaned from the test itself about the environment (Ideal)
  - Information you have been provided by the target organization (Secondary Option)

A Simple Evaluation Model for Pen Test Maturity

A Simple Model to Evaluate Pen Tests

- There are a lot of ways to evaluate pen tests
- A simple model for maturity and assessing a given test includes 3 categories:
  - **Methodology**: Aspects of the testing regimen itself, including tools and tactics
  - **Realism**: How the test reflects real-world attack vectors and attacker strategies and focal areas
  - **Reporting/Output**: Contextual and informative output
- Each category has 5 key points to evaluate

Category: Methodology

- Does a stated methodology exist, or is an industry methodology being employed? (Yes/No)
- Are all major phases of testing incorporated, from Reconnaissance to Exploitation? (Yes/No)
- Are both manual testing and automated tools involved? (Yes/No)
- Is actual exploitation allowed? (Yes/No)
- Is pivoting from one compromised system to another permitted? (Yes/No)

Methodology Scoring

- Each Yes=1, each No=0
- 1-2: The methodology/approach in use is likely immature and will miss many potential vulnerabilities
- 3-4: The methodology is in line with “standard” pen testing practices, with room to improve/enhance
- 5: The methodology is advanced and represents a mature pen testing approach tactically
- **Note: This is a **very** subjective category, and will depend on testers, tools, etc.

Category: Realism

- Is a “black box” approach taken? Does the tester have little knowledge of the target environment? (Yes/No)
- Were efforts made to avoid detection? (Yes/No)
- Was “malware” dropped or emulated in some way? (Yes/No)
- Was social engineering involved/included? (Yes/No)
- Was data (or “mock” data) exfiltrated? (Yes/No)

Realism Scoring

- Each Yes=1, each No=0
- 1-2: The pen test is likely somewhat unrealistic relative to actual attack scenarios seen in the wild today
- 3-4: The pen test uses tactics commonly seen in real intrusion scenarios
- 5: The pen test was planned to emulate real-world threats and attack vectors in numerous ways
- **Note: This category is less applicable for “SOP” and “Point and Shoot” test types

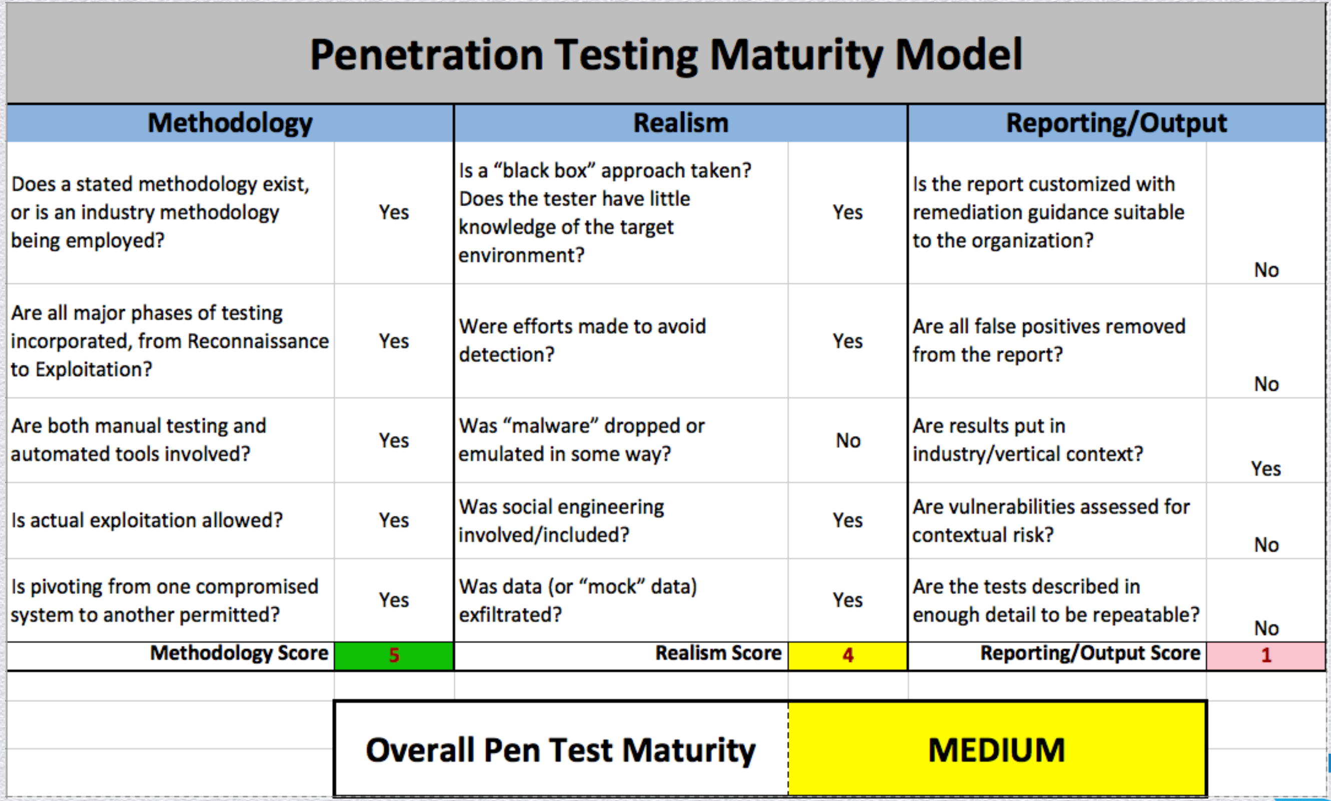
Category: Reporting/Output

- Is the report customized with remediation guidance suitable to the organization? (Yes/No)
- Are all false positives removed from the report? (Yes/No)
- Are results put in industry/vertical context? (Yes/No)
- Are vulnerabilities assessed for contextual risk? (Yes/No)
- Are the tests described in enough detail to be repeatable? (Yes/No)

Reporting/Output Scoring

- Each Yes=1, each No=0
- 1-2: The report/output is mediocre at best, and the pen test likely provides little value
- 3-4: The results are likely valuable to the organization, but could still be more tailored and useful to multiple stakeholders
- 5: The results are well-crafted and provide high business and operational value

The Model In Total



Penetration Testing Maturity Model

Methodology		Realism		Reporting/Output	
Does a stated methodology exist, or is an industry methodology being employed?	Yes	Is a "black box" approach taken? Does the tester have little knowledge of the target environment?	Yes	Is the report customized with remediation guidance suitable to the organization?	No
Are all major phases of testing incorporated, from Reconnaissance to Exploitation?	Yes	Were efforts made to avoid detection?	Yes	Are all false positives removed from the report?	No
Are both manual testing and automated tools involved?	Yes	Was "malware" dropped or emulated in some way?	No	Are results put in industry/vertical context?	Yes
Is actual exploitation allowed?	Yes	Was social engineering involved/included?	Yes	Are vulnerabilities assessed for contextual risk?	No
Is pivoting from one compromised system to another permitted?	Yes	Was data (or "mock" data) exfiltrated?	Yes	Are the tests described in enough detail to be repeatable?	No
Methodology Score	5	**Realism Score**	4	**Reporting/Output Score**	1

Overall Pen Test Maturity	**MEDIUM**

Download here: www.voodoosec.com/PenTestMaturityModel.xlsx

Conclusion

- Pen tests have been around for 40+ years, but still aren’t really “mature”
    - Many disagree on format, true value, etc.
- There’s a LOT we can do to improve
    - These are just a few suggestions based on my experience as a practitioner and consultant
- This maturity model is just a starting point, and can readily be improved and built upon. Let’s do this!

Questions?