Cybersecurity Metrics: A Red Team Perspective

Bradley R. Horton
CISSP-ISSMP, CEH, CISA
Chief, TSMO Red Team
24 Apr 2015

Author Info

- Supported hundreds of cyber assessments including physical security, OPSEC, and social exploitation
- Prior life with non-DoD public sector and commercial securitywithin the private sector
- Leads a group of ‘ethical hackers’ and security assessment professionals charged with exploiting hardened and security systems and commands within the DoD
- Holds several industry certs and is a proud graduate of the University of Alabama

Agenda

- What is a Red Team?
- Subjectivity versus Objectivity in Cyber
- Cyber attacks are not binary bullets
- Suggestions from a Red Team (Bad, Better, Best)
- Questions

Red Team

- Blue, Red, Purple, White, Green, even Pink
  - Confusion of the terms can lead to retesting
  - All have a role to play
  - The Army and “TCNO”
- Accreditation and Authorities for DoD Testing
- Expected results and findings differ significantly
- Testing goals and philosophies differ significantly
  - Threat emulation versus Penetration Testing
- Support to system acquisition and development should not differ
  - We’re all in this together (Find ->Fix -> Verify)

Subjectivity and Objectivity

- How do we measure cyber success when:
  - System characteristics differ wildly between, e.g. business system versus C2 system versus weapon systems
  - Assessor (Red Teams or the like) capabilities vary
  - Threat definitions are in abstract and nebulous terms
  - Cyber threats can develop instantly versus a slow evolutionary development

Cyber Attacks are Not Binary Bullets

- Bad choices
  - "Number of Attacks Attempted"
  - "Number of Successful Attacks"
  - "Number of Attacks Detected"
  - "Number of Attacks Prevented"

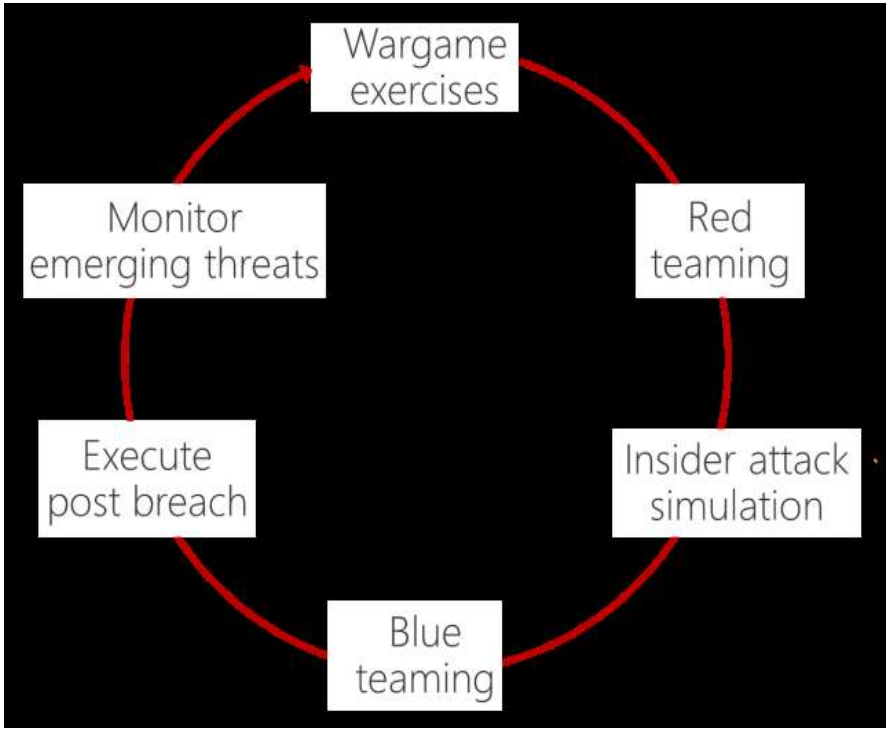
  Discussion: On the surface these sound great. If the Threat launches 20 attacks, that's better than 10. If 2 out of 10 are successful, that's better than 5 (for the SUT). If the system prevents 9/10 attacks, that's better than 6/10. If all 99% of all attacks are detected, that's better than 50%.

  Answer: No. It is 100% dependent on protecting and defending what matters.

  The fundamental difference, again, between a real threat and a penetration test: Real threats only have to find one way to achieve their goal. If the goal is achieved…

Suggestions from a Red Team

- Better
  - Microsoft Model – white paper on Red Teaming the Enterprise Cloud
- Mean Time To:
  - Compromise
  - Escalate
  - Detect
  - Recover
- Sounds similar to the “PDRR”
- Discussion:
  - Allows comparisons and provides some level of objectivity
  - Flaw: Different teams with different skills? Different operators of defenses?



Cyber Attacks are Not Binary Bullets

- Best(ish)
  - Incorporate measures of team and tool sophistication
    - Red Teams can objectively measure the complexity of exploitation
      - » System “Hygiene” [default creds?]
      - » Known exploit [MS08-067?]
      - » Common vulnerability [SQL Injection?]
      - » Open Source Vulnerability [Field manual with password]
      - » Chained exploit requiring several combinations of attacks
    - Red Team can objectively define methods to remain hidden
      - » Implant sophistication
      - » Web-shells and encrypted C2
      - » Disk Residency or in-memory ‘tricks’
    - Red Team can subjectively define expected impacts
      - » Total control
      - » Total disruption

Cyber Attacks are Not Binary Bullets

- Best(ish) continued
  - Measures of React/Restore > measures of Protect/Detect
  - Restated: Resiliency > Security

  - React/Restore Metrics:
    - » Can the system enable operators to detect, react, and restore within a mission window
    - » Missile System Example

Closing Thoughts/Questions

- All ‘color’ teams have their purpose; understanding the differences is key
- Subjectivity is still rampant in cyber assessment – qualitative versus quantitative
- Metrics which do not measure the complexity of attack or the impact of the attack are often meaningless
- Measures of React/Restore > measures of Protect/Detect
  Restated: Resiliency > Security
- Knowing 999,999 attacks failed and only one succeeded does not necessarily indicate success

Questions