5 years in adversary emulation

Does Threat Intelligence have a valid role in testing security resilience?

James Chappell – Co-Founder and Chief Innovation Officer
@jimmychappell
In 25 minutes

- Adversary Emulation: brief history
- Experience with CBEST
- Update on TIBER
- Key Takeaways
- The Future?
- Was it worth it?
For this presentation:

- I do not represent or speak on behalf of CREST, The Bank of England, Financial Conduct Authority, DNB, ECB or any other regulatory institution – I am simply sharing publicly stated learnings from experience.

- I am not able or willing to share details of specific tests but will talk in general about experiences from them.

- Digital Shadows do not currently offer CBEST, or TIBER (EU/NL) tests but may do in the future – a good thing: means I can be super honest and direct about our experiences without fear of harming future businesses.

- Journalists – please make yourselves known, hopefully this is more about where we take the profession overall, but if you want to write about this I can help!
A journey
5 (and a bit) years

**Blue Teams**
- Concept of CBEST announced by BoE
- Tests underway

**Red Teams**
- Preparation
- Execution

- **2014**
- **2015**
- **2016**
- **2017**
- **2018**
- **2019**
- **Future?**

---

- **Nov ‘17 - Dutch National Bank – TIBER – NL**
- **May ‘18 Tiber-EU announced**
- **Oct ‘18 GBEST (UK Gov)**

---

**Also in UK**
- Telecoms
- Nuclear
- Aviation

---

**Monetary Authority of Singapore (MAS)**

**Hong Kong Monetary Authority**
Phase I - CBEST
In May 2014, the Bank of England along with the professional body CREST launched CBEST and STAR testing frameworks.

CBEST introduced a threat led approach to conducting security testing.

Goals:
1. Realistic tests based upon a set of evidence of threats observed in the wild. Tailored to the customer.
2. Hold institutions accountable to testing being a true test of resilience.
3. Broader in scope than a traditional pen test (a red team approach) focused on critical economic functions (CEF).
Drivers: Professional and skilled Red Teams are important but...

- Sometimes solely focused on technical outcomes with technical stakeholders - struggle to involve business stakeholders but “managed by IT/InfoSec team”
- Follows well trodden paths (for good reason, but not articulated why)
- Often conducted work separately from organizations risk assessment
- Regulators want to hold institutions to account to justify tests are true measures of resilience rather than tech for tech sake
- Regulators want boards to get involved in their managing their risks
- Testing often change driven with scope set by what is new, rather than what is important

NOTE: Intelligence should be a way of *supporting* a Red Team not dictating actions.
Why do intelligence before a red team at all?

Tests focus on the PROBABLE threats rather than the theoretically POSSIBLE
### Threat Intel in CBEST: Key outputs

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Goals</th>
<th>Evidence</th>
</tr>
</thead>
</table>
| - Threat scenario  
  - Based on detailed research  
  - Emulating real threat  
  - Tailored to **Target** assets | - A set of Goals for the test team  
  - A set of agreed ‘flags’ the team must capture | - A **lot** of Supporting Evidence to show that the test is real  
  - Validated by UK Gov |

**SUPPORTS SELECTION OF TARGET and TEST PLAN**

**PRIORITISES “FLAGS” AGAINST GOALS AND MOTIVATION**

**BACKS UP BUSINESS CASE FOR MITIGATING CONTROLS**
Model Overview

Output: Threat Scenarios to be used in a test
Threat Intelligence Products

1. Threat Intelligence Report

- Provides analysis of threat groups based on thorough research
- Evidence to justify and support actions of testing team
- **OUTPUT**: Threat Scenarios
- **USE CASE**: Provides supporting evidence for use in security test.

2. Targeting (Foot printing) Report

- Broad analysis of digital footprint to identify riskier areas
- **NOT** a full reconnaissance exercise
- **OUTPUT**: Initial targets for test
- **USE CASE**: Provides input into reconnaissance phase of security test.
## Threat landscape

<table>
<thead>
<tr>
<th>Threat source</th>
<th>Capability</th>
<th>Intent/activity</th>
<th>Threat score to Client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insider intentional*</td>
<td>H</td>
<td>H</td>
<td>16</td>
</tr>
<tr>
<td>Nation State – Disruption and Attack (CNA)</td>
<td>VH</td>
<td>M</td>
<td>15</td>
</tr>
<tr>
<td>Nation State – Espionage (CNE)</td>
<td>VH</td>
<td>M</td>
<td>15</td>
</tr>
<tr>
<td>Organised Crime – Economic</td>
<td>H</td>
<td>M</td>
<td>12</td>
</tr>
<tr>
<td>Nation State Proxy</td>
<td>M</td>
<td>M</td>
<td>9</td>
</tr>
<tr>
<td>Hacktivist</td>
<td>L</td>
<td>M</td>
<td>6</td>
</tr>
<tr>
<td>Journalist/researcher</td>
<td>L</td>
<td>L</td>
<td>4</td>
</tr>
<tr>
<td>Organised Crime – Extortion</td>
<td>M</td>
<td>VL</td>
<td>3</td>
</tr>
<tr>
<td>Insider unintentional</td>
<td>VL</td>
<td>VL</td>
<td>1</td>
</tr>
</tbody>
</table>

Scoring based on high watermark assessment
CBEST intelligence and testing processes

威胁情报提供者

威胁情报
- 获得可信赖的当前威胁情况视图

目标
- 模拟敌对方的战略

活动和规划
- 根据可靠证据制定安全测试计划

安全测试
- 执行红队安全测试

规划和风险评估

威胁情报报告
- 威胁摘要
- 威胁概要
- 威胁剧本

目标报告
- 人类目标
- 过程目标
- 系统目标

目标活动
- 红队安全测试规范

测试报告
- 红队测试结果
THREAT PROFILES CONSIDERED

- FUZZYSNUGGLYDUCK
- APT7334
- FANCYMOOSE
- Angry Cyber fighters (CNA)
- AnonUnChuffed
Threat Scenarios follow a narrative structure
Mapping to a storyline
Mapping to a storyline

- **Reconnaissance**
  - Tools, tactics, techniques, and procedures
- **Preparation**
- **Infiltration**
- **Entrenchment**
- **Compromise**
- **Exploitation**
- **Impact on target**
- **Post-attack TA actions**

- **Rising action**
- **Climax**
- **Falling action**
- **Dénouement**

**Key**
- Threat intelligence
- Scenario-specific additions
CBEST - What Went Well

- Created an evidence backed business case for a broad end to end test of resilience/red team where hard to justify previously
- Created useful discussion on what is ‘critical & economically important’ separate from tech change.
- Forced organizations to prove IR playbooks were really working to regulators
- Genuinely got the board to take the test seriously and helped understand the challenges
- Created discussion about what is probable and linked to other risk assessment
- Took business stakeholders end to end through process helping to justify existing investments in defenses and Detection and Response capabilities
<table>
<thead>
<tr>
<th>Observation</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Bank X and National Bank Y have pretty much the same threats –</td>
<td>Shared threat models better where this is shared - but ”opportunities” for attackers different due to varying tech stack – need a</td>
</tr>
<tr>
<td>often a validation of what was already known</td>
<td>common threat model and shared labour. Also only true for sub-types. Infrastructure, Investment Banking vs. Retail Banking.</td>
</tr>
<tr>
<td>The Red Team still carried out the same test</td>
<td>Not intended to dictate red team, but help justify actions.</td>
</tr>
<tr>
<td>The scenarios would benefit from being more specific</td>
<td>Tools such as MITRE ATT&amp;CK give us increased specificity now we would have benefited from that then</td>
</tr>
<tr>
<td>It was labour intensive</td>
<td>Yes – components should be made generic and shared x-industry where possible.</td>
</tr>
<tr>
<td>After the Red Team made initial intrusion discoveries were made that did</td>
<td>Yes – should be an interactive continuous process</td>
</tr>
<tr>
<td>not relate to the scenario</td>
<td></td>
</tr>
<tr>
<td>After initial intrusion scenarios written in absence of internal recon</td>
<td>Both Scenarios and test plans should only be finalized after initial intrusion.</td>
</tr>
<tr>
<td>needed updating</td>
<td></td>
</tr>
</tbody>
</table>
Phase II – The TIBER(s)
• Progressive approach – learnings from tests quickly integrated into approach and standards
• Created a shared ‘Threat Landscape’ document on which tailored threat scenarios can be developed, greatly reducing the labour required during the threat phase – more cost effective
• Better handover and collaboration between threat intelligence and testing provider updating test plans and scenarios in light of findings during test
Summing up – Where should this go?*

* In my humble opinion
MAKE IT PURPLE
Combining outputs

Red Team
- Testing plan
- TTP’s
- Attacker Objectives

Blue Team
- Incident Response Playbooks
- Detection Strategies
- Control Configuration

Continuous Validation
- Operationalize
- Attack/Defense
- Agile Response Planning
Biggest takeaways

• **Make it Purple**: Instead of passing threat reports over, continually update and validate throughout the test. Make Red Teams inform the Blue Team and vice versa. Make it a continual test of the IR playbooks, make regulatory test a snapshot of this embedded process.

• **Operationalize this**: Threats change constantly – Should not be a one-off test: Embed threat modelling into Incident Response, and Preparedness planning on a continuous basis – demonstrate on ongoing basis and then pick examples once a year.

• **Involve the business throughout**: No better model of a threat than an incident (a threat/risk that came to pass). Businesses know their critical assets from an internal perspective better than anyone – this is all valid input.

• **MITRE ATT&CK Adversary Emulation Plans** – A threat model with real purpose and community collaboration, A common language for Threat Intelligence and Red Teams to talk to each other but also increasing utility across the board

• **Share and Share-a-like**: Shared Threat Landscapes and Efficient Collaboration tailoring for just the efficient.
The Future

• Automation in Vulnerability Management – Platforms such as ATTACKIQ, SafeBreach etc taking real scenarios and including them in routine testing

• MITRE ATT&CK provides a very helpful model which should exist throughout these tests and be the center for them, adversary emulation.

• Pen Testing Frameworks:
  • Cobalt Strike (C2 emulation and in memory artefacts)
  • Caldera (open source framework)
  • APT Simulator
  • Metta
  • Blue Team Training Toolkit (BT3)

Great resource list here: http://pentestit.com/adversary-emulation-tools-list/
Does Threat Intelligence have a valid role in testing security resilience?

YES

- A justification for a broad test
- A live measurement of the ‘playbook’ in realistic circumstances
- A way of ‘trying out’ threat intelligence, or comparing it to existing feeds or capability
- Validation of existing thinking and controls, risk and response plans
- Evidence to support business cases

Use a regulatory driver to support a business case – to achieve the things you wanted to do anyway