

SIGDEV Conference 2012

(U) Making Things Measureable: Technology Trending Challenges and Approaches

June 2012

Derived from NSA/CSSM 1-52
Dated 20070108
Declassify On: 20370501



Overview (U)

- (U) Setting the Stage
 - Strategic Surprise, Priority Needs, Definitions

- (U) Making Things Measurable
 - Emerging Technology Discovery
 - Technology Use Discovery

- (U) Challenges
 - Complexity
 - Getting data is only step 1
 - Visualization
 - Building outreach and engagement



CT Trends Focus Questions (U)

(U) Does NSA CT know what technologies, communications products and applications, and modus operandi *are being used* by terrorists, terrorist groups, or in locations of interest?

(U) Does NSA CT know what emerging technologies, communications products and applications, and modus operandi *are likely to be used* by terrorists, terrorist groups, or in locations of interest?

***Prevent Strategic
Surprise***



CT Trends Focus Questions (U)

(U) Does NSA CT know what technologies, communications products and applications, and modus operandi *are being used* by terrorists, terrorist groups, or in locations of interest?

(U) Does NSA CT know what emerging technologies, communications products and applications, and modus operandi *are likely to be used* by terrorists, terrorist groups, or in locations of interest?

(C//REL) What we're really asking is:
Can we tell which ones are likely to become a priority need?



Risk Management for SIGINT Threats (U)

- (S//REL) Threat to SIGINT Capability
 - A behavior or technology that has the potential to have a *negative impact* on NSA's capability to provide SIGINT on a Terrorism Target
- (U) Use Risk
 - The possibility that a particular threat will be adopted by Terrorist targets
- (S//REL) Indications and Warning
 - Early warning of high impact threats to prevent surprise to key stakeholders and reduce risk from Terrorist adoption of technology that would adversely affect SIGINT production

(S//REL) NSA's ability to manage risk is directly proportional to our ability to detect threats



The data-driven approach (U)

“Count what is countable, measure what is measurable, and make measurable that which cannot be measured”

Galileo (17th century astronomer)

“When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind”

Lord Kelvin (discovered absolute zero)

“You cannot manage what you cannot measure”

Bill Hewlett (co-founder of Hewlett-Packard)

“Not everything that counts can be counted, and not everything that can be counted counts”

- Albert Einstein



So... what is a (CT) trend? (U)

A trend is a **measurement** of occurrence

(S//REL) Comparing the behavior of a *single* target...

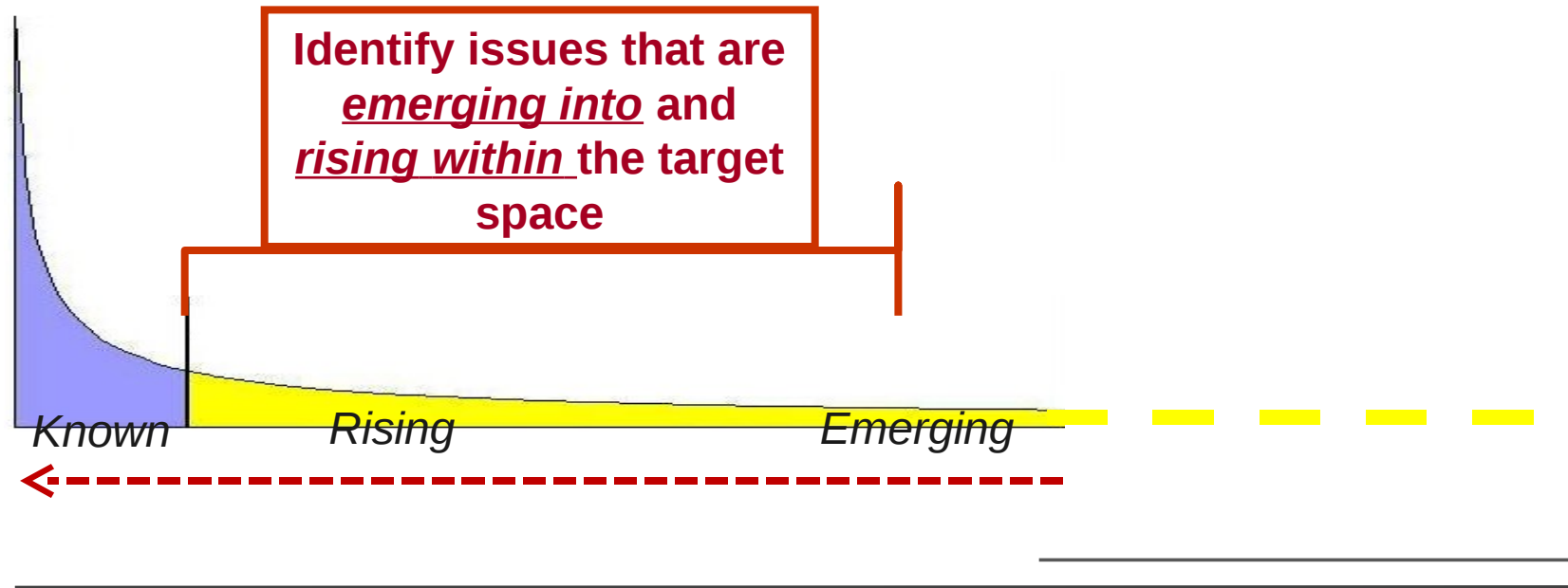
- Pattern-of-life
- Modus Operandi
- Technology Usage

...to the behaviors seen within the *target space*

- Multiple targets, ***within and across the entire CT enterprise***
- Over a period of time



Prediction and Identification of Priority Needs Prevents Strategic Surprise (U)



Making Things Measurable (U)

Emerging
Technologies

Technical
Thought Leaders

Technologies
in Use



Innovation Phases (U)

Adoption

Experimentation

Interest



Technology Adoption Factors

(U)

• Availability

• Vested

• Authority

• Trusted Source

• Availability of



Optics (U)

- (S//REL) Optic #1: Emerging Technology Discovery
 - Focused primarily on interest and experimentation phases of innovation
 - Watching the Watchers
 - Weaker indicators
 - ***New technologies***

- (S//REL) Optic #2: Technology Use Discovery
 - Focused primarily on adoption phase of innovation
 - Owning the Known
 - Stronger indicators
 - ***New targets***



Analytics and Processes (U)

Emerging	Both	Technology
Producti on	Producti on	Producti on
Technical Thought Leader Scanning	Technical Thought Leader Scanning	Technical Thought Leader Scanning
Production Element Scanning Note Project	Production Element Scanning Note Project	Production Element Scanning Note Project
FORTREND - Extremist Technical Sub-Forum	FORTREND - Extremist Technical Sub-Forum	FORTREND - Extremist Technical Sub-Forum
Administrative Response Technology Use Discovery	Administrative Response Technology Use Discovery	Administrative Response Technology Use Discovery
Forensic	Forensic	Forensic
Seized Media Trends Tracking	Seized Media Trends Tracking	Seized Media Trends Tracking
CNE Trend Tracking and Discovery	CNE Trend Tracking and Discovery	CNE Trend Tracking and Discovery
Network	Network	Network
User	User	User
Top X - Active	Top X - Active	Top X - Active
Technology Pattern of Life Analytic (Automated)	Technology Pattern of Life Analytic (Automated)	Technology Pattern of Life Analytic (Automated)
Mobile Handset	Mobile Handset	Mobile Handset



Optic #1: Emerging Technology Discovery (U)

- (S//SI//REL) Emerging Technology & Behavior Discovery
 - Detection of *interest, experimentation, knowledge transfer* or *direction* using content, metrics approaches
 - Currently using deskside & virtual engagement to leverage TOPI analyst initiative to discover, prioritize, and work against “**strongest**” indicators
 - Leverages inherent TOPI expertise and functions of traffic processing/translation/tasking etc..
 - Embedded analysts, virtual relationships: production “customers”
 - Currently identifying, tracking ‘technical’ thought leaders
 - Technical sub-forums, scanning notes measurements
 - Administrative emails (No-Reply etc..)
 - Forum links, uploaded/downloaded files

Goal: Generate Prioritized Input (techs/behaviors) for Research



Optic #2: Technology Use Discovery (U)

- . (S//SI//REL) Technology Use and Behavior Discovery
 - _ “Stratactical” data sets
 - _ Includes target-specific data point for each item (e.g. selector)

 - _ Discovery of target behavior by identifying technology use patterns, trends, and/or anomalies in:
 - _ User-agents (browsers, OS, devices)
 - _ Tasking (new tasking, total tasking)
 - _ Network, Protocol usage (Active User metrics)
 - _ Visited URLs, web searches
 - _ Process lists, pre-fetch logs, registry entries, software logs
 - _ Hardware usage (smartphones, tablets, SD cards)

 - _ Currently using various tools (XKEYSCORE, SEEKER, BIONICTURTLE, JEMA, JOLLYROGER, MARINA, TUNINGFORK, QFDs, etc...) and approaches with multiple cloud analytics in varying stages of development and/or planning

Goal: Generate Prioritized Input (techs/behaviors) for Research



Measurement Drives Research (U)

(S//REL) Triage begins with **target** indicators of a new technology

Derived from either optic: Emerging or Use Discovery

Interest, Experimentation, Use, Knowledge Transfer, Metric, etc...

Target \Rightarrow Technology \Rightarrow Do other targets use this technology?

This is the central defining question for Trends Analysis:

Do other CT targets use this technology?



Weak vs. Strong Indicators: Brutal Triage (U)

Weak	Medium	Strong
• E	• E	• E
• I	• I	• I
• No	• No	• No
• Received a	• Received a	• Received a
•	•	•
• Exper	• Exper	• Exper
• Previous/Low	• Previous/Low	• Previous/Low
• Installed, no	• Installed, no	• Installed, no
•	•	•
• A	• A	• A
• Hig	• Hig	• Hig
• Log files, traffic	• Log files, traffic	• Log files, traffic



The Wicked Problem Aspect (U)

(S//REL) Defining the problem is the first (wicked) problem

- Triage Stage 1
- Initial priority: (single) target + initial understanding of technology
- Implications Research
- What does the product/service do?
- Current NSA capabilities to detect, collect, exploit, analyze?
- Do any other CT targets use it?
- Triage Stage 2
- Updated priority: target(s) + updated understanding of tech/USSS
- Validated Next Steps
- As needed: capabilities/access development requirements
- Reporting: internal, CIR, e-gram; Gaps report; prioritization w/in tech category



Goal: Periodic Reporting Vehicle (U)

- . (U//FOUO) Move beyond ad hoc task responses to routine deliverables
- . (U//FOUO) Overcoming volume challenge
 - Huge variety of inputs, massive numbers in each
 - Prioritization
 - Visualization
- . (S//REL) Moving threats to a simple Risk Assessment model
 - Borrows methodology from models used for executive purposes elsewhere in agency
 - (FAMT, Geopolitical Technology Trends Matrix, TAO...)
 - Opportunities, threats handled separately



Capabilities Development Risk

Matrix (II)

<p>Impact</p> <p>></p> <p>to production</p> <p>Use Risk</p> <p>v</p>	<p><u>TRIVIAL</u></p> <p>Loss/lack of insight to small aspect of target communications, presence</p>	<p><u>MINOR</u></p> <p>Loss/lack of insight to significant aspect of target communications, presence</p>	<p><u>MODERATE</u></p> <p>Loss/lack of insight to large component of target communications, presence</p>	<p><u>MAJOR</u></p> <p>Loss/lack of insight to majority of target communications, presence</p>	<p><u>CATASTROPHIC</u></p> <p>Near-total loss/lack of insight to target communications, presence</p>
<p>Current Highest Priority Target Use</p>	<p>Document tracking</p>	<p>Fivewes, Facebook chat presentation</p>	<p>Mail.ru, TeamViewer, Join.me</p>	<p>OTR, Tor, Smartphones, Zoho.com webmail, TrueCrypt</p>	<p>Tor+ Trilight Zone + Cspace + ZRTP VoIP client on Linux</p>
<p>Current Operational Target Use</p>					
<p>Current Low Priority/Previous Higher Priority Target Use</p>					
<p>Technical Thought Leader Recommendations, Experimentation</p>					

Capabilities Development

Impact > to production Use Risk V	<u>TRIVIAL</u> Loss/lack of insight to small aspect of target communications, presence	<u>MINOR</u> Loss/lack of insight to significant aspect of target communications, presence	<u>MODERATE</u> Loss/lack of insight to large component of target communications, presence	<u>MAJOR</u> Loss/lack of insight to majority of target communications, presence	<u>CATASTROPHIC</u> Near-total loss/lack of insight to target communications, presence
Current Highest Priority Target Use	Document tracking	Fivewes, Facebook chat presentation	Mail.ru, TeamViewer, Join.me	OTR, Tor, Smartphones, Fastmail, TrueCrypt	Tor+ Trilight Zone + Cspace + ZRTP VoIP client on Linux
Current Operational Target Use					
Current Low Priority/Previous Higher Priority Target Use					
Technical Thought Leader Recommendations, Experimentation					

(TS//SI//REL) With rare exceptions, application-specific solutions are only built based on these two criteria

Capability Development Challenges (U)

(TS//SI//REL) With rare exceptions, application-specific solutions are only built based on these two criteria????

- In resource-restrained environment, development of capabilities against ***likely-to-increase in priority*** applications is trumped by standing requirements driven by ***known priority*** applications
- Capabilities development response to current/priority technology threats occurs normally w/in existing resources - but response does not scale, either to the industry or to multiple crises



Simplifying the Risk Matrix (U)

Impact > to production Use Risk <	<u>TRIVIAL</u> Loss/lack of insight to small aspect of target communications, presence	<u>MINOR</u> Loss/lack of insight to significant aspect of target communications, presence	<u>MODERATE</u> Loss/lack of insight to large component of target communications, presence	<u>MAJOR</u> Loss/lack of insight to majority of target communications, presence	<u>CATASTROPHIC</u> Near-total loss/lack of insight to target communications, presence
Current Highest Priority Target Use	2nd Quadrant		1st Quadrant		
Current Operational Target Use					
Current Low Priority/Previous Higher Priority Target Use	4th Quadrant		3rd Quadrant		
Technical Thought Leader Recommendations, Experimentation					

Adding in the Solution

<p>Impact > to stakeholders Use Risk v</p>	<p><u>TRIVIAL</u> Automated solutions not required; manual workflows sufficient: XKS fingerprint</p>	<p><u>MINOR</u> Minimum automated solutions required: realm creation, simple extraction, presentation</p>	<p><u>MODERATE</u> Significant, routine capdev required: STARPROC capability, CES detectors, endpoint characterization, correlation analytics</p>	<p><u>MAJOR</u> Requires focused inter-office capdev: Lead/Program manager needed, SPF required, inter-office/2p/3p working group</p>	<p><u>CATASTROPHIC</u> Requires SID-level attention: FAMT, LAE; WPMO, AMOD portfolio integration</p>
<p>Current Highest Priority Target Use</p>	<p>2nd Quadrant</p>		<p>1st Quadrant</p>		
<p>Current Operational Target Use</p>					
<p>Current Low Priority/Previous Higher Priority Target Use</p>	<p>4th Quadrant</p>		<p>3rd Quadrant</p>		
<p>Technical Thought Leader Recommendations, Experimentation</p>					

Examples: Jan-February 2012

(TS//SI//RFI)

Impact > to production Use Risk v	<u>TRIVIAL</u>	<u>MINOR</u>	<u>MODERATE</u>	<u>MAJOR</u>	<u>CATASTROPHIC</u>
	Loss/lack of insight to small aspect of target communications, presence	Loss/lack of insight to significant aspect of target communication s, presence	Loss/lack of insight to large component of target communications, presence	Loss/lack of insight to majority of target communications, presence	Near-total loss/lack of insight to target communications, presence
Current Highest Priority Target Use	TeamViewer			Tor	
Current Operational Target Use	Join.Me			TrueCrypt	
Current Low Priority/Previous Higher Priority Target Use	LaplinkGold			TAILS	
Technical Thought Leader Recommendations, Experimentation	Muslima				
	Purematrimony.com			Web.de	
	Zemana Anti-Keylogger			Cspace	
				Redphone	

Goal: Emerging Technology Snapshot (U)

- (U) Executive version – snapshot of top items only
- (S//REL) Overcoming the challenges of prioritization and volume is still only 50% of the problem
- (S//REL) Stated Preference:
 - Breakdowns by target/target set
 - Preserve opportunity vs. threat
 - Identify HUMINT sources for collaboration



Emerging Technology Snapshot (U)

Target/Org	Tech	Quadrant
AQSL courier	TAILS	1
GIMF	TAILS	
AQ media	TrueCrypt	
S2I42	Join.Me	2
LT, S2I42	TeamViewer	
LT	Laplink	
TTL	Extremist version of Tor	Opportunity
AQ media	Encrypted Webmail	Source

(TS//SI//REL) Full details available as needed



Emerging Technology Snapshot (U)

- (S//REL) Monthly Emerging Technology Snapshot
 - 1-3 page Snapshot (6 page max if previous month data included) to CT leadership
 - Snapshot + supporting full data to MICROEXPANSE
 - Underlying processes in alpha stage
 - Stopgap until maturation of multiple efforts
 - Data Explorer, ECHOBASE
 - Inclusion of FAA/PRISM in GM-Halo



End Results – Tactical & Strategic (U)

- . (S//REL) Tactical Outcomes
 - **Lead Generation**
 - Target Development
 - Target Discovery
 - Behavior Detection
 - Access Prioritization

- . (S//REL) Strategic Outcomes
 - Prioritization for **Capabilities Development**
 - Driven by target priority: single target + volume of targets
 - Prioritized within tech category, target (set) category
 - Overall CT product line prioritization



Challenges (U)

- . (C//REL) Complexity
 - **Understand target, technology, & SIGINT system**
- . (S//SI//REL) Getting data is only step 1
 - **Getting a data set is like to getting a new bearer to analyze**
- . (U) Visualization
 - **Excel tops out at a million rows...**
- . (TS//SI//REL) Clean data
 - **Targets vs. Selectors**



Overcoming Complexity (U)

SIGINT System SIGINT System

Fingerspitzengefühl

- *Literally “fingertip feeling”*
- *Empathy, sensitivity, tact*
- *Ability of military commanders to react rapidly*

CT Trends Team: Technology Technology

SIGDEV analysts

Partner/Enablers

Must understand tech threat implications, provenance and structure of data to manipulate, interpret it



Getting Data is Step 1 (U)



Getting Data is Step 1 (U)

**Every Step Takes Time, Effort, Tools
and Most Importantly: People to Do
the Work**



Getting Data is Step 1 (U)

Be customer driven. The ability to get data is endless. The ability to do work isn't.



Visualization (U)

(TS//SI//REL) Excel tops out at a million rows...

- 19 branches, 30+ target sets, ~200 realms, ~800 domains, ~45000 selectors = *1 million rows/~2.5 weeks* for summarized active user events from EO12333 alone
- Spreadsheets are good, but not everyone knows how to use a pivot table
- Each dataset can easily provide 4-5 or more pivoted looks for each branch/target set = *minimum 100-150 slides*

(S//REL) Intent is to routinely produce multiple large datasets on a monthly basis for collection management, research purposes



Visualization (U)

(S//REL) Analysts work at the selector level

- Leadership wants data presented at the target level

(S//REL) Automated population of technology, behavior information in analyst workflow tools, databases

(S//REL) Each separate visualization task takes manpower, time away from operational analysis



Clean Data (U)

- (S//SI//REL) Metrics will only provide a near-accurate picture: ground truth will always be the domain of the TOPI and based on content
- (S//SI//REL) Some selectors (accurately) map to multiple targets, multiple teams, multiple organizations
- (S//SI//REL) Some selectors simply don't have a known target, only a target set
- (S//REL) Need to correlate across widely different datasets requires creation of normalized bridge datasets (e.g. comparing executables to domains)
- (S//SI//REL) TKB/UTT are victims of years of "fill in the blank" freeform data entry; very slowly being addressed (~2015?)



Rising Strategic Issues (U)

- (TS//SI//REL) Encrypted Webmail Services
 - Atabmail, Zoho, Safe-mail, Fastmail, HMA Mail
- (TS//SI//REL) Remote Desktop Viewers/Remote Access Tools
 - TeamViewer, Join.me, Cybergate
- (TS//SI//REL) Aggregators/Over-the-Top Messaging Services
 - WhatsApp, Nimbuzz, eBuddy



What Next? (U)

- . (S//SI//REL) Continue to build, strengthen, expand:
 - internal workflows, research and discovery capabilities

 - collaboration with production elements
 - Operational support via embedded analysts at NSAW
 - Tradecraft, technical support virtually with extended enterprise

 - partnerships with FVEY SIGDEV community
 - Establish and expand dialogue opportunities
 - “Failure Sharing” - tradecraft sharing and operational deconfliction

- . (S//REL) Technology Trends MyNoc



fin

Question
s?



Comment
s?

