Identifier Lead Triage with ECHOBASE

NSA - S2I51
NSA - T1442

JUN 2012
The Problem

SIGINT is very good at 2 things:

1. Establishing lists of potential leads (50-10k+)
2. Manual analysis to vet individual targets
A common model for identifier lead lists, today:

Input: Seed List Provided to SIGDEV

Phase 2: Normalize and Expand Selectors

Phase 3: Foreignness and Compliance Check

Phase 4: SIGINT Queries on Selector activity and behavior attributes

Bulk enrichment of ‘SIGINT business knowledge’

Manual analysis
After initial enrichment checks, the analyst is often left with too many identifiers of “possible interest”

- Possible Interest (Pri 2): 75%
- Definite Interest (Pri 1): 5%
- No Further Analysis Needed: 20%

Percentages are conceptual
Bulk Lead Triage via Behavior Analytics

- Hundreds or thousands of selectors to go through high level vetting very quickly
- Better triage prioritization allows for highly adjustable thresholds to be set for follow-on analysis
- Compliance can be inserted at both the “batch result” and “query” level
- Potentially utilize multiple clouds & cross-enterprise analytics

![Pie chart showing interest levels:]
- No Further Analysis Needed: 20%
- Low Interest (Pri 4): 25%
- Medium Interest (Pri 3): 35%
- High Interest (Pri 2): 15%
- Definite Interest (Pri 1): 5%
Identifier ‘SIGINT Business’ Enrichment

Bulk gathering, via Identifier Scoreboard (phase 2/phase 3)

- Targeting
- Authorities
- Reporting
- Targets
- Knowledge
- Foreignness
- Compliance

...not a raw SIGINT query
‘Yes/No’ Identifier Behavior

Bulk triage, via SIGINT Analytics Mode (start of phase 4)

Core set of ‘yes/no’ behavioral questions about a set of identifier leads

...against raw SIGINT!
SIGINT Analytics Mode

Triage by aggregate behaviors

Quickly zero in on worthy leads
SIGINT Analytics Mode – Detailed View
### SIGINT Analytics Mode – Detailed View

**Had direct communications with a targeted identifier?**

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012-Mar-04 09:58:46 Z</td>
<td>was bcc’d on email from [redacted]</td>
<td>UUID: [redacted] SIGAD/PDDG: DS-200A / C4 Case notation: [redacted] Legal authority category: EO12333</td>
</tr>
</tbody>
</table>

**External links to guide next steps in analysis**

- Go view target knowledge
- Go view content
- Add new knowledge
Initial set of analytic questions
- Most running within GHOSTMACHINE framework
  - Limited contributors
- GHOSTMACHINE Analytic Engine provides
  - QFD hosting of analytic results
  - RESTful query interface

Future analytics
- multiple organizations/frameworks
2012 Olympics Sharing

Job Tracker

GCHQ

Seeded Mic
Seeded Analytic
Seeded Mic
Seeded Analytic

(GCHQ architecture details omitted)

NSA

User DN, justification, leads & which QFDs ("domains")

Lineup query details

User DN, justification, leads & which QFDs ("domains")

Log queries

Check user authorizations

Check user authorizations

WAVELEGAL

CASport

GM Analytic Engine

Releasable targeted identifiers

QFDs Svc

OCTAVE

UTT

Targeting

Targeted identifiers

Selector List

Seeded Mic
Seeded Analytic

Non-GM Analytic

Analytic

Bulk feed of analytic results

Bulk feed of analytics results

GHOSTMACHINE

TOC SECRET//COMINT//REL TO USA, CAN, AUS, GBR, NZL
2012 Olympics Support

• NSA SID Leads Evaluation Cell
  • Triage of Olympics-based leads through the event
  • Leverage both NSA and GCHQ-produced analytics

• Greater SID-wide usage following the Olympic period
Contact/Information

- Briefers:
  -
  -
- ECHOBASE Alias:
  -
- NSA WikiInfo page:
  -