Discussions in Deep, Dark, and Open: Monitoring the Social Discourse of Cyberattack

Oct 10, 2013

Jack Zaientz, Bob Bechtel, Matt Hollingsworth
Jzaientz@soartech.com
bob.bechtel@soartech.com

SOARTECH

Modeling human reasoning.
Enhancing human performance.

SOARTECH

Support cyberattack analysis through social discourse

- **What Needs to Be Done:** Provide cyber analysts with better tools for monitoring large social media streams by
  - detecting expected and unknown items of interest
  - detecting deliberate attempts to spread disinformation, incite hostilities, propagate cyberattack tools
  - assessing source quality
- **What's Hard:** We have a haystack, but are not sure that it's needles we're looking for
  - Difficult to detect unknown facts (new cyber-attack methods)
  - Difficult to manually develop filters that capture the complexity of the domain
  - Difficult to understand map of community boundaries relative to cyberattack development
- **What we're doing */ Core Insight*:**
  - Developing a search query grammar and associated filter algorithms and visualization tools to ***leverage the innate linguistic and communication structure of social media***
- **Timeline:**
  - **Have done initial experimentation under IRAD funds. Will be developing prototype interfaces and algorithms over next two years with goal of operational testing**

Insight/hypothesis: Much as frequency analysis is a core cryptanalysis technique because it exploits known regularities in language, social discourse analysis can be a core technique in social media analysis because it exploits known regularities in language (including communications patterns, term use)

SOARTECH

Social Media as OSINT, both HUMINT and COMINT

- HUMINT – *direct observation, where sensor is secondary to observation*
  - Intentional or inadvertent "civilian sensors" with eyes on a situation
  - Comparing pre- and post- event social media behavior to gauge civilian reaction and attitudes
  - Disaster Impact and Battle Damage Assessment imagery

New Checking for the new SabPub malware in OS X http://t.co/wU

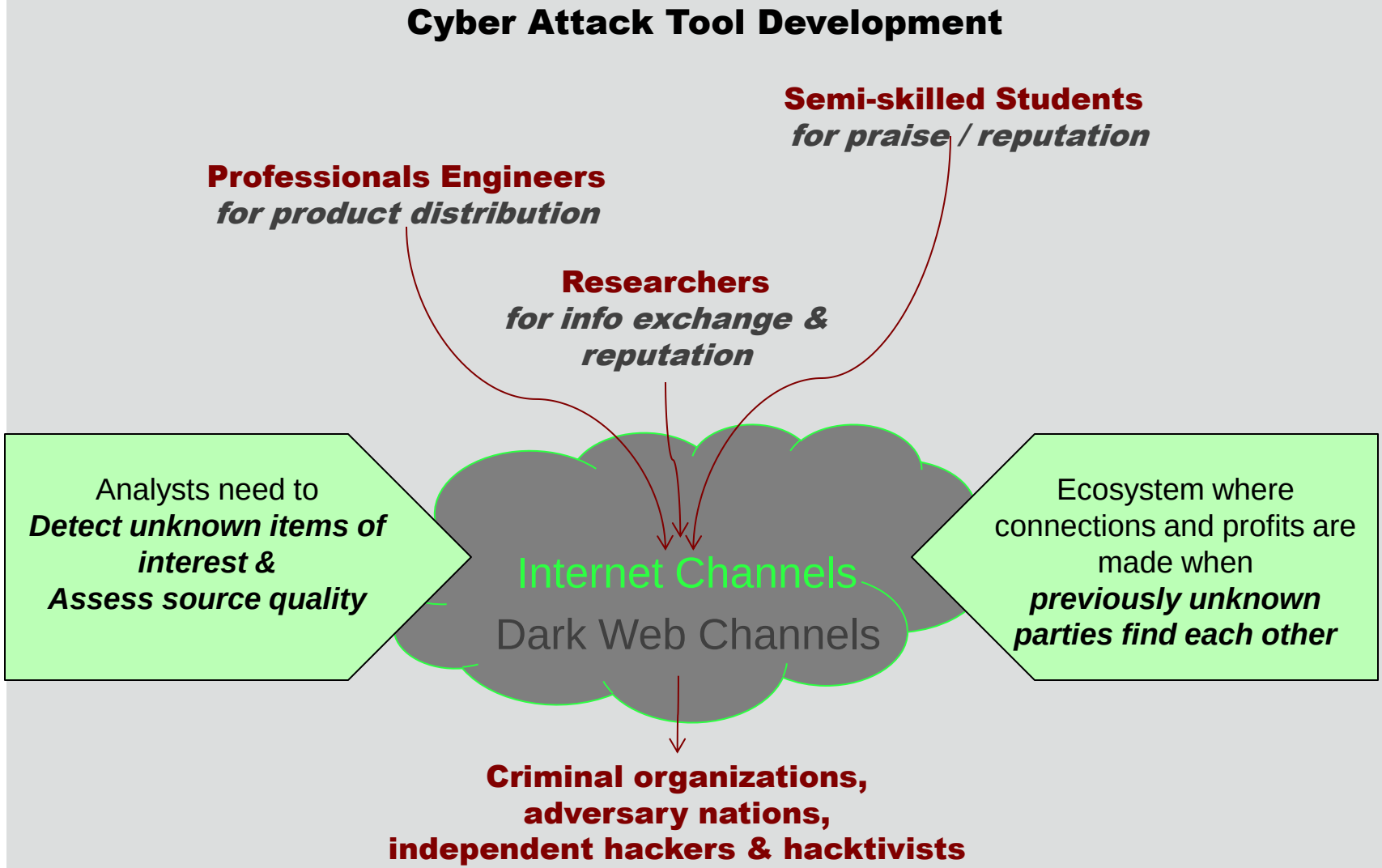
new malware version

- COMINT – *intercepted communications, where sender/receiver is critical*
  - Adversarial influence operations & disinformation
  - Requests for support (HA/DR supplies, intervention)
  - Advertisement / movement of contraband (particularly cyber)
  - Claims of responsibility for attacks
  - Development of trans-national networks

New Checking for the new SabPub malware in OS X http://t.co/wU
#occupycanada #occupytoronto #occupyvancouver
#occupymiami #occupydc

"political hacktivists" are aware of new malware version

CyberAttack communication in the open web



The internet is used by to facilitate communication between anonymous or previously unknown-to-each-other collaborators

Monitoring social media for *communication*

- **Leveraging social media communication structure helps overcome big-data scaling issues**
  - but new end-user tools are needed…
- *Community-specific language:*
  - *Communities use specific terminology that can be used for filtering, but it's challenging to discover and results in combinatorics in filter definition*
- *Communication Patterns:*
  - *Communities use stable communication patterns ; e.g. novelty signaling, shouting/#beaconing*

Novelty Signaling

Combine novelty & cyberattack terms

Novelty: new, novel, recent
CyberAttack: access control, adware, arbitrary memory, bot, botnet, brute force, buffer overflow, buffer size validation, cache poisoning, ddos …

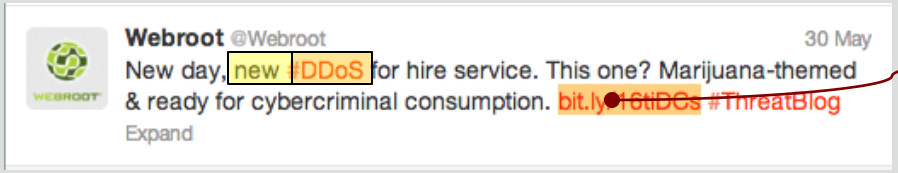
Recent experiment used over 300 term pairs to filter 1 year of twitter fire-hose to less than 5 cyber-relative tweets per day.

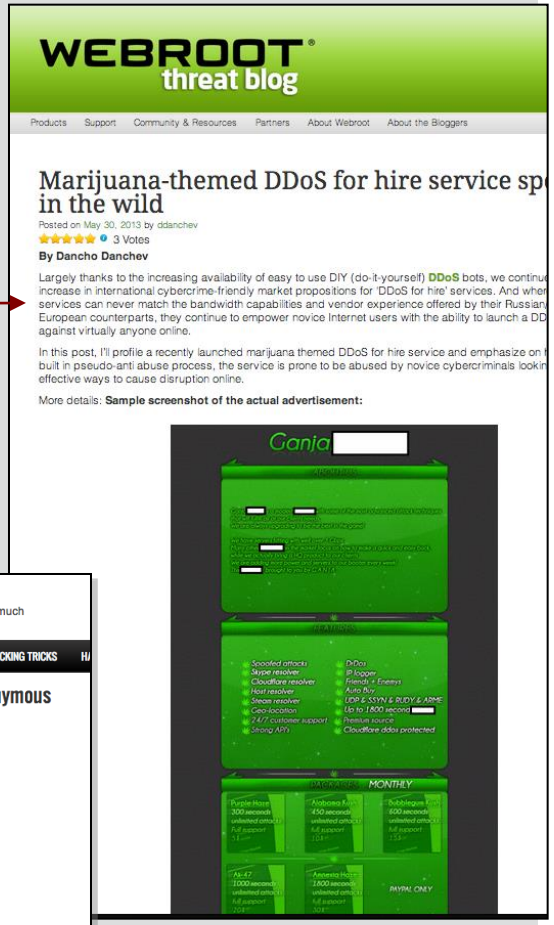
Twitter Shout

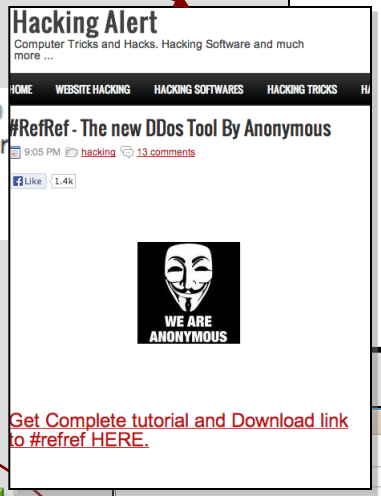
Using multiple hashtags to make a tweet findable by target communities

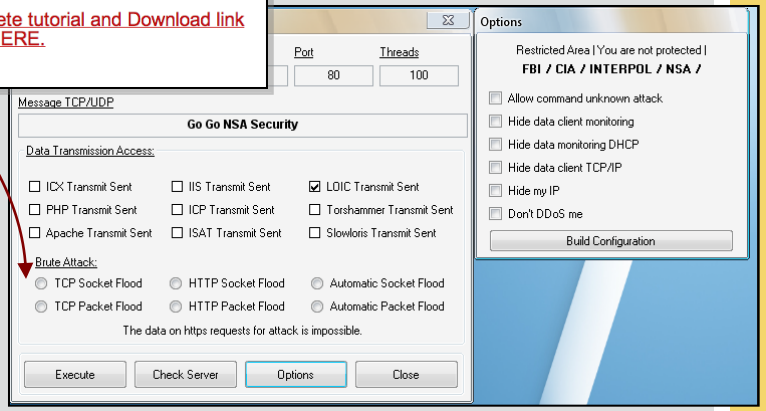
Adversaries, collaborators, and at-risk populations want to be found.
But we need to be listening properly

More examples from the wild…..









Novelty Signaling behavior is common… but analysts can't handle vast combinatorics with text-literal query interfaces

Next Steps for Cyber and Crisis analysis SA

- On the ground now
  - Improve understanding of communication flows and community vocabularies within target domains to anticipate deliberate communication
  - Use current keyword based (e.g. ONRs TweetTracker) tools in a systematic manner to develop term use base-lines and to monitor for key tweets
- Future S&T
  - Develop next-generation analysis tools that ease burden on analysts and improve search results by leveraging social discourse structure

Are you a hashtag? If I tweeted this, would it reach you?

Important workshop on social media impacts on human security & military operations. Be there. #nato #onr #dhs #red cross #UNHCR #GDACC #social media #humansecurity #cybersecurity #resiliency #disasterrelief #civilconflict #crisisoperations #privacy

SOARTECH

Modeling human reasoning.
Enhancing human performance.