Traffic Analysis
The Most Powerful and Least Understood Attack Methods

Raven Alder, Riccardo Bettati, Jon Callas, Nick Matthewson
What is Traffic Analysis?

• Signals intelligence that ignores content
• Information for analysis is the metadata

• “Traffic analysis, not cryptanalysis, is the backbone of communications intelligence.”

Susan Landau and Whitfield Diffie.
Interesting Metadata

- Endpoint addresses
- Timing
  - Duration
  - Sequencing
- Location?
- etc.
Why is it important?

• The title of the panel says it all
  – We are going to startle you
• Everyone needs to think differently
  – Often we’re protecting the wrong thing
  – TA shows the limits of possible defense
• Potential for new research and creativity
Historic Uses

- Finding size, scope, intentions of military
- Marketing research
- Reconfigure networks
Why do this?

• Crypto
• Too much data, already
• It’s easier than analyzing everything
• It’s hard to defend against
Defenses

• Include
  – Don’t communicate
  – Don’t be seen communicating
    • Spread Spectrum, etc.
  – Insert false communications

• Naïve defenses often worse than nothing
  – Everything you know about this is wrong
What can we do?

- Determine alert status of military
  - Notorius “Domino’s Metric”
- Identify authors of text
  - “Primary Colors” break against Joe Klein
- Crack SSH passwords
  - *Timing Analysis of Keystrokes and Timing Attacks on SSH* [Usenix 2001]
What more can we do?

- Identify OS of remote hosts
- Identify host as it moves around the net
- Correlate virtual-to-physical hosts
- Unravel mix nextworks
What else can we do?

- Remove text redacting
  - [http://cryptome.org/cia-decrypt.htm](http://cryptome.org/cia-decrypt.htm)
- Identify movies being played
- Identify music being downloaded / played
  - CDDB finds albums with TA-like methods
And even more

- De-multiplex IPsec tunnels
- Spatially locate hosts
- Voice analysis of some speech patterns
- Analysis of social networks
- Google PageRank
- nmap, p0f
- Credit card fraud detection
Open Questions?

- How do we guard against TA?
- How do we use TA?
  - Can it be used against spam, botnets?
  - Are there offensive and defensive uses?
Additional Reading

- “Introducing Traffic Analysis” by George Danezis
  - [http://one.revver.com/watch/147903](http://one.revver.com/watch/147903)