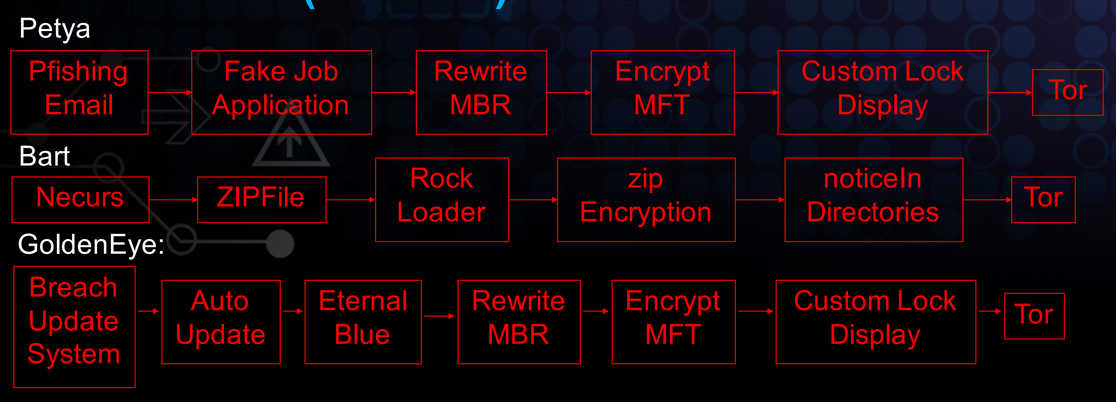
Ransomware Analysis as Dialog for Attribution and Reconnaissance (RADAR)

H. Van Dyke Parunak, PhD, ABC Research, LLC

Physics | Linguistics | Biology

ABC Research, LLC

A Tale of Two (or more) Attacks:



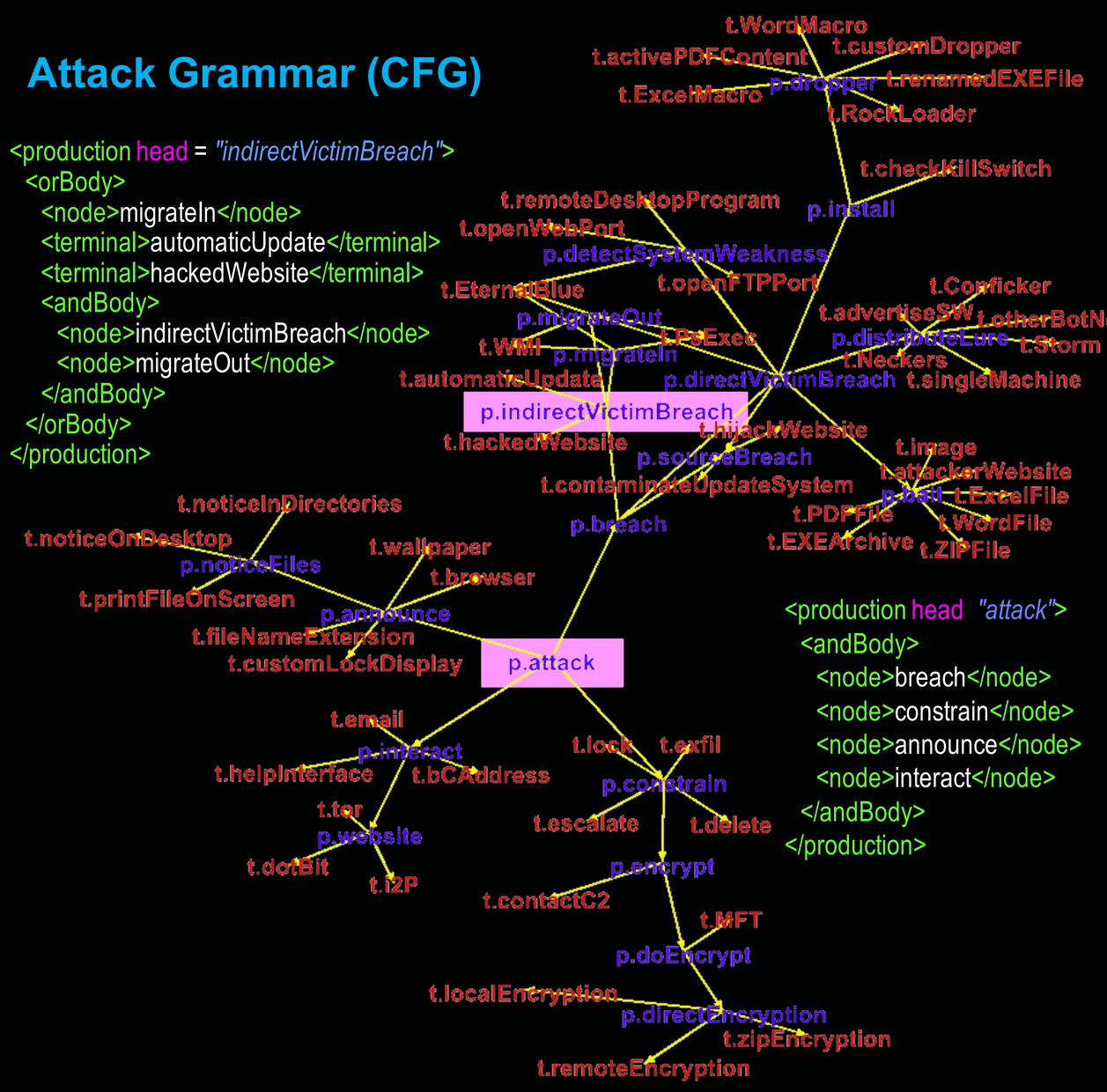
1. To whom should we attribute each attack?
2. How consistent is joint attribution of two attacks?

Initial Data:

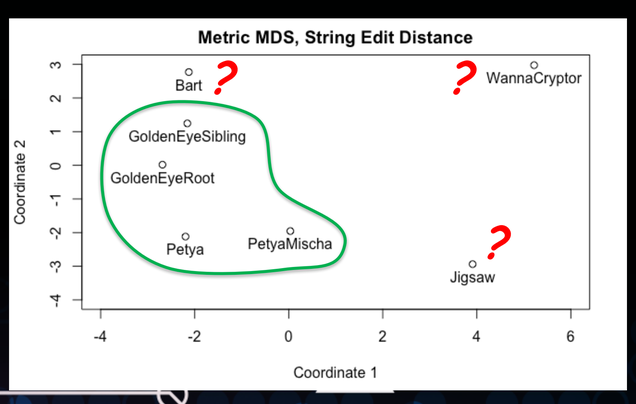
Attack	Appeared	Example Distinctives
Bart	June 2016	Local encryption via zip files
GoldenEye initial	Jan 2017	Distribution via fake job application
GoldenEye derivative	June 2017	Distribution via SW update
Jigsaw	April 2016	Incrementally deletes files if ransom not paid
Petya	April 2016	Encryption of master file table rather than files
Petya with Mischa	May 2016	Petya with fall-back conventional encryption
WannaCry	May 2017	Breach via Eternal Blue NSA exploit; kill switch

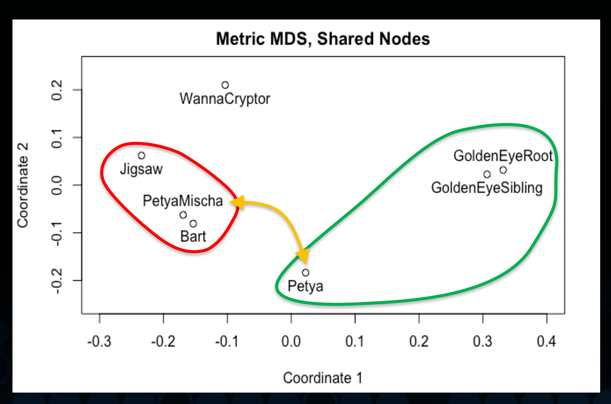
Attack Grammar (CFG)

```
<production head = "indirectVictimBreach">
 <orBody>
  <node>migrateIn</node>
  <terminal>automaticUpdate</terminal>
  <terminal>hackedWebsite</terminal>
  <andBody>
   <node>indirectVictimBreach</node>
   <node>migrateOut</node>
  </andBody>
 </orBody>
</production>
```




```
<production head "attack">
 <andBody>
  <node>breach</node>
  <node>constrain</node>
  <node>announce</node>
  <node>interact</node>
 </andBody>
</production>
```





Key thesis of RADAR:

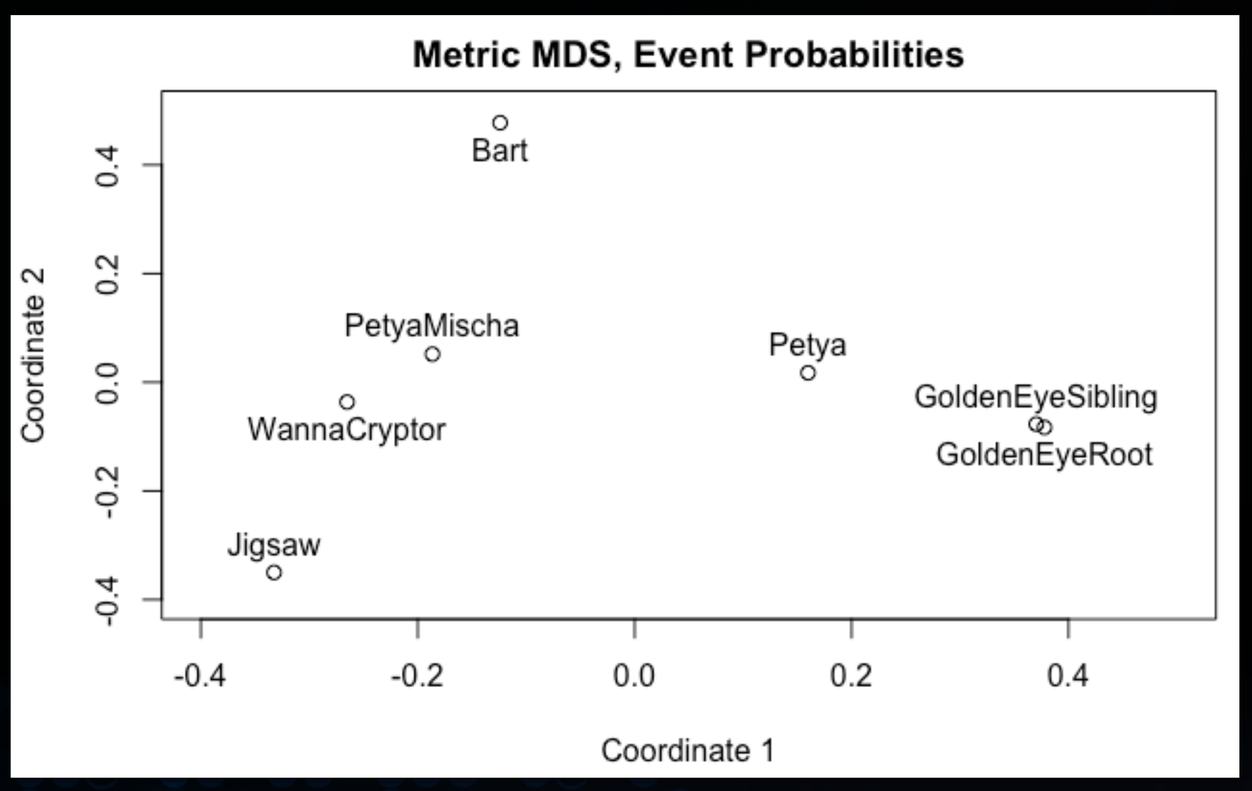
1. Current methods of attribution are based on *isolated* characteristics of an attack.
2. Ransomware involves the victim in a *dialog* with the attacker.
3. This dialog can be characterized *linguistically* to identify organic patterns.
4. These patterns *integrate details* to help attribute attacks.

Universal Attack Structure

	Attacker / Computer	Third Party	Victim's Computer	Victim	Stage
1	→	→	→		Breach System
2			↻		Apply coercion
3	→	→	→	→	Announce attack
4	←	Bitcoin ←	←	→	Obtain payment
5	→	→	→		Release

Proximity Measures on Attack Sequences:

- Baseline: string edit distance (no grammatical info)
- Grammar-based:
  - Via Lempel-Ziv compression (cf. DNA comparison): needs repeated substrings
  - Shared nodes: assumes balanced grammar
  - Probabilistic: compare event probabilities with probabilities conditioned on shared prefixes



SCORE
SPECIAL CYBER OPERATIONS RESEARCH AND ENGINEERING

C3E