

Malware Phylogenetics

Penny Chase

Desiree Beck

781-271-2113 • pc@mitre.org

714-915-0310 • dbeck@mitre.org

MITRE Sponsored Research

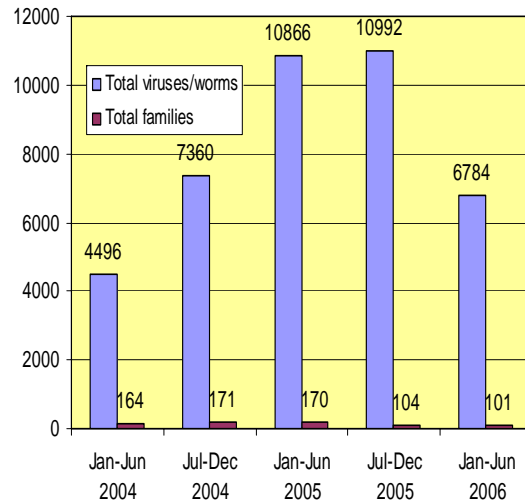


Problem

- **Understand the evolution and relationships between malware threats**
 - **Anti-malware industry and end users need to know whether anti-virus products offer protection and how to remediate.**
 - **Cybercrime investigators need to know if a new threat is related to an ongoing investigation and need to develop leads towards attribution.**
 - **Malware researchers/analysts want to leverage previous analyses.**

Background

Rise of Malware Variants



One of the major factors contributing to the increase in previously unseen threats is the number of variants within malicious code families... attackers are commonly updating current malicious code to create new variants.

Symantec Internet Security Threat Report, September 2006

The nature of threats is changing from widespread to targeted and regional... [Malware creators] are motivated more than ever by financial gain.

Trend Micro, *The Trend of Threats Today: 2006 Annual Roundup and 2007 Forecast*



MITRE

© 2007, The MITRE Corporation

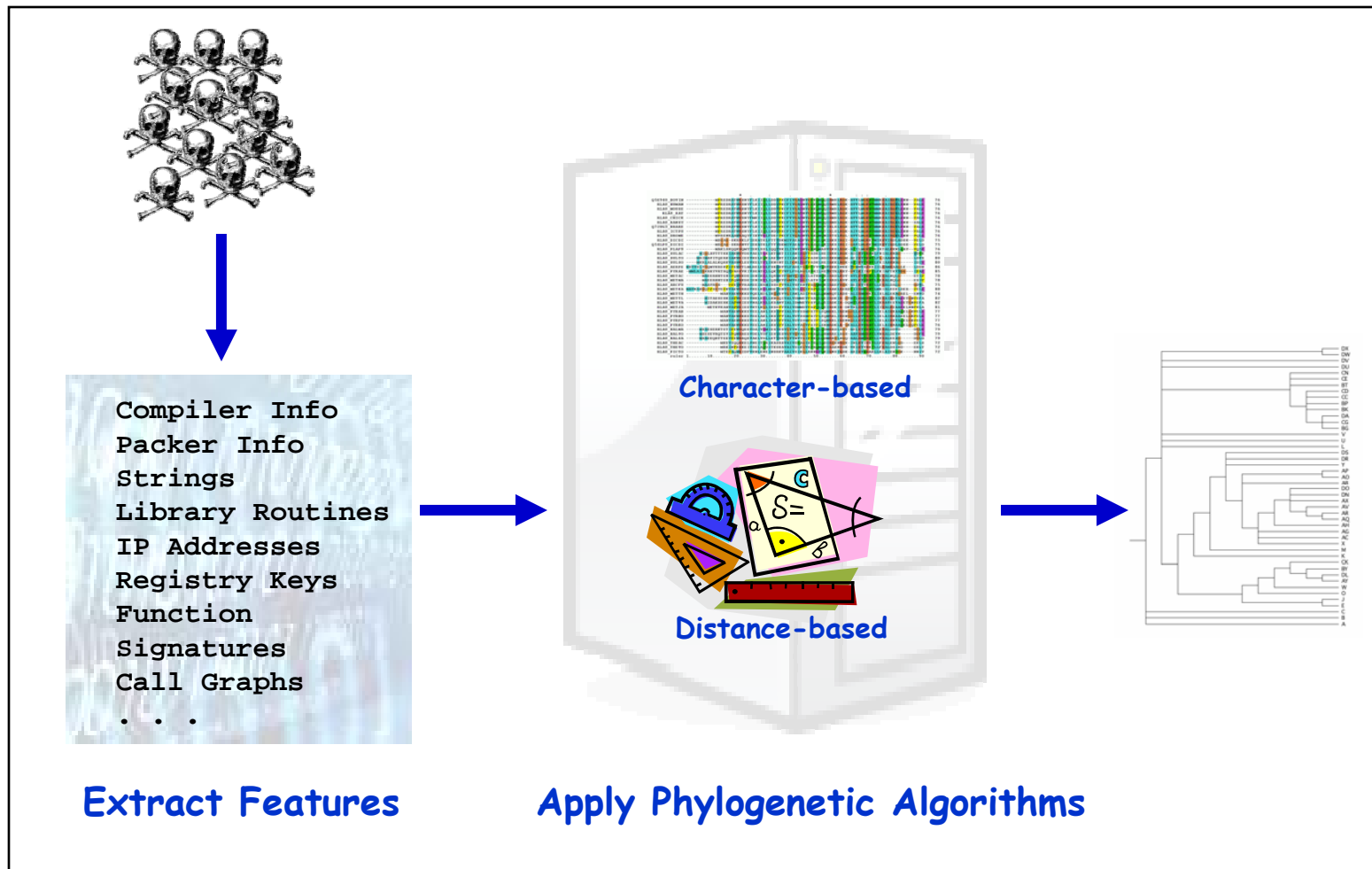
Objective

- **Apply the bioinformatics algorithms used by biologists to infer phylogenetic models to malware analysis in order to reveal the evolutionary relationships between malware threats**
 - **Create phylogenies for specific malware**
 - **Develop criteria for good phylogenetic modeling features**

Activities

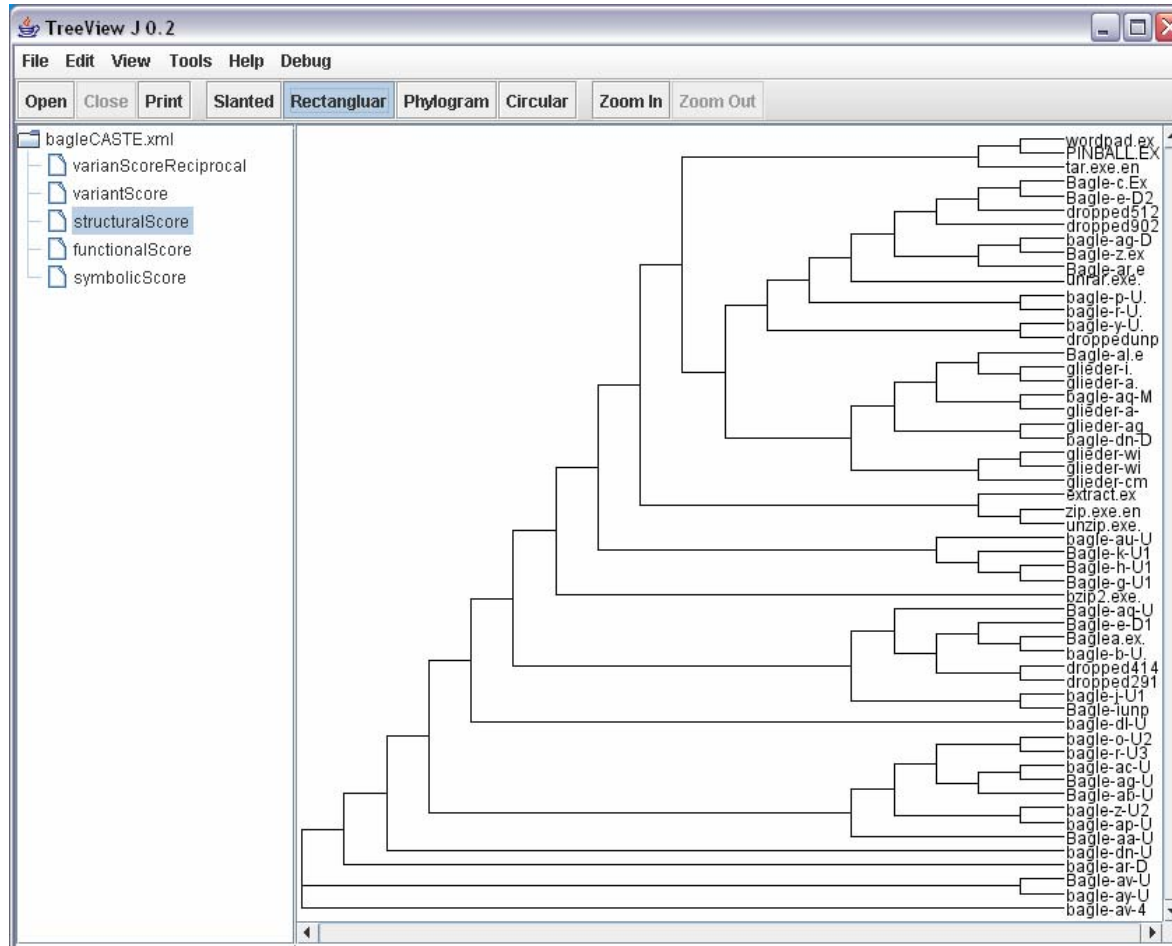
- **Build data set of unobfuscated malware samples**
- **Develop “ground truth”**
 - **Manual analysis of samples**
 - **External information (e.g., research paper, anti-virus vendor encyclopedias)**
- **Select and extract features**
- **Create phylogenetic models**
- **Run experiments using different features and algorithms**

Highlight



Generate Phylogenetic Models of Malware

Demonstration



Phylogenetic Tree of Bagle Variants

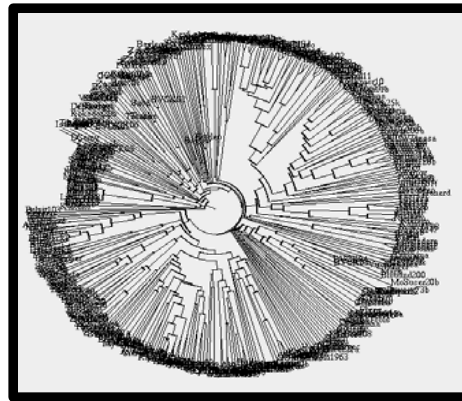
Impacts

- **Malware researchers/analysts: improve collection-based malware analysis**
- **Cybercrime investigators: insight into malware authorship and potential leads for attribution**
- **Anti-malware industry and end-users: better detection, prevention, remediation, and classification of malware**
 - **DHS/DoD Software Assurance Forum
Malware Working Group**

Future Plans



Experiment with other features (e.g., align code segments using BLAST)



Develop phylogenies for multiple malware families

Explore code reuse (true lineages vs common libraries)

