Attribution has been studied for various purposes, including its use in trying to attribute attacks to the attackers that carry them out. In one review of the issues [7.13] the following characterization was provided:

Differentiation By How We Use Symbols To Communicate

N-Gram Analysis and Other Statistical Methods

Unigrams, Bigrams and Trigrams

Tokenization

Unigrams

Bigrams

- Dice Coefficient
- Fishers exact test - left sided
- Fishers exact test - right sided
- Log-likelihood ratio
- Mutual Information
- Pointwise Mutual Information
- Odds Ratio
- Phi Coefficient
- T-score
- Pearson's Chi Squared Test

Trigrams

- X-Gram: Largest Token Size of Value
- Hidden Markov Models
- Morphemes and Phonemes
- Gender Identification
- Authorship Attribution
- Unrelated But Applicable Bag-Of-Word Techniques
- Non Bag-Of-Word Similarity Techniques

Differentiation by How We Type

- Original IBM-Selectric Work
- Biometric Authentication Work
- SSH Timing Attack Work
- Unrelated But Applicable Biometrics Techniques

Differentiation by How We Attack

- Log File Analysis
- Attack Tree Comparison
- Attack Code Similarities
- Attack Code Eggs That Uniquely Identify the Attacker

Basic Premise

Attackers are created over time

- We acquire from the media we utilize
- The skills that are learned matters
- You can only use tricks you know
- The tricks have an originator and a path of dissemination
- Trails of access are not hidden farther back in the attack

The time that a skill is used matters

Copyright © 2003-2009 Chet Uber & Fred Cohen. All Rights Reserved.

Newbies fumble around in the dark
Hackers make syntax mistakes
Gurus make few mistakes and enter complex commands

Attacker Profiling
Types of Attackers
Building the Reference Base

Pattern Recognition

Representation
Extraction
Classification and Identification

Ways to Characterize an Attacker

Biometrics

Keystroke Dynamics

Not what, but How
Habitual Typing Rhythms

Keystroke Verification

Static
Continuous

Stylometrics

Coding Style

HTML from hack sites
HTML from USENET groups
Code Samples from all books on programming
Code from information security sources
Code from cryptography sources
Code from hacking sources

Graphic Design Style
Vocabulary
Behavioral DNA
n-gram Analysis

[7.13] C. Uber , Personal correspondence, Dec. 2003 - Feb. 2004.

Copyright © 2003-2009 Chet Uber & Fred Cohen. All Rights Reserved.