ON THE DESIGN AND EXECUTION OF CYBER-SECURITY USER STUDIES: METHODOLOGY, CHALLENGES, AND LESSONS LEARNED

Malek Ben Salem & Salvatore J. Stolfo

CSET 2011

INTRODUCTION

- Problem
  - Lack of masquerader data
  - Schonlau data set not appropriate
- Objectives
  - Test the conjecture that extensive search reveals an attacker’s malicious intent
  - Evaluate whether decoy files embedded in a local file system can be used to detect masqueraders
- Steps
  - Conduct user studies to validate conjecture
  - Gather new dataset including data from “normal users” and masqueraders

USER STUDY METHODOLOGY

- State hypotheses
  - Experimental hypothesis
  - Null hypothesis
- Identify experimental variables
  - Independent variable
  - Dependent variables
  - Confounding variables
- Build control groups
  - Scenario narratives to control user’s intent
- Determine sampling procedure
- Estimate sample size
  - Power analysis

USER STUDY EXECUTION

- Obtain IRB approval early
- Develop/deploy the right sensors for data collection
  - Right unique IDs
  - Right platform
- Pilot experiment
- Reduce confounds and bias
- Sanitize data and have users sign waivers
- Validate collected data after reviewing post-experiment questionnaires

HYPOTHESES

- Experimental hypothesis
  - If the intent of a masquerader is malicious, then they will engage in a ***significant*** search activity on the victim's system.
- Null hypothesis
  - The manipulation of the masquerader's intent does not have any significant effect on the masquerader's search behavior.

→The observed significant effect during the experiment can be attributed to the manipulation of the masquerader's intent and cannot be the result of pure chance.

IDENTIFY EXPERIMENTAL VARIABLES

- Independent variable
  - Only variable manipulated by researcher, all others are kept constant
  - Need one control group for each value of the independent variable
  - User's intent
- Dependent variables
  - Observed behavioral feature to be measured by researcher
  - Tightly dependent on independent variable
  - User's search behavior
- Confounding variables
  - Random variables affecting observed behavioral feature
  - Need to be eliminated or at least minimized
  - E.g. Awareness of monitoring , familiarity with desktop search tools

BUILD CONTROL GROUPS

- Control user's intent through scenario narratives
  - One narrative for each control group
  - Milgram's experiment
- Scenario narrative requirements
  - Generalizable: representative of masquerade attack
  - Conforming to threat model
    - Assumptions should be clearly stated
  - Detailed
    - Includes answers to anticipated questions to limit verbal communication with study participants
    - Minimizes user bias
  - Easily executable
    - E.g. time-limited

SCENARIO NARRATIVES

- User has access to coworker's system for 15 minutes while coworker is away
- Malicious, benign, and neutral scenarios

Experimental Variable	**Value**	**Same/ Different**
Scope	Local File System of Colleague's Computer	Same
Environmental Constraints	IDS Lab Computer	Same
Desktop Configuration	Same Recent Documents and Applications	Same
Time Constraints	15 minutes	Same
Intent	**Malicious, Benign, or Neutral**	**Different**

DETERMINE SAMPLING PROCEDURE

- Objective: Increase the sensitivity of the experiment
- Means: Reduce uncontrolled variability
- Subject variability makes up the largest source of variability
- Sampling procedures
  - Use same subject in all treatment conditions
    - Violates assumption in our threat model that attacker is not familiar with victim’s file system
  - Use homogeneous group of subjects
    - Similar characteristics relevant to experiment
  - Use several small subject sets
    - Sets highly homogeneous within one set, but widely varying between sets

PERFORM POWER ANALYSIS

- Power
  - Indicates how statistically significant experiment's results are
  - Desirable values: 0.5-0.9
  - Used to determine required sample size for each treatment condition
  - Higher power requires larger samples
- Adequate sample size* depends on
  - Number of independent variable and number of treatment conditions
  - Desired effect size that researchers wishes to detect
  - Desired power

*KEPPEL, G. Design and analysis : a researcher's handbook. Pearson Prentice Hall, 2004.

REDUCE CONFOUNDS AND BIAS

- Reduce subject variability
  - Homogeneous group of user study participants
- Reduce experimental treatment variability
  - Same desktop for all experiments
  - Same file system contents: automated collected data upload
  - Same recent documents opened for each participant
  - Same researcher

USER STUDY EXECUTION

- Obtain IRB approval
  - Lengthy, iterative process
  - Required very detailed information
    - e.g. call for participation, data items collected
- Develop/deploy sensors for data collection
  - Study technology market trends to select the right development platform
- Pilot experiment
  - Learn sources of variability
- Sanitize data
  - Data collected for same user from different sensor s
  - Users did not take advantage of sanitization functions provided
- Review post-experiment questionnaires
  - Extract trends, eliminate invalid cases

RUU (ARE YOU YOU?) DATASET

- Characteristics
  - Larger than 10GBytes in size
  - More tan 10 million records
  - Data from 18 "normal" users
    - 4 days of data on average
  - Data from 40 "masqueraders"
- Results
  - Search behavior reveals malicious intent
  - Search behavior profiling detects100% of masquerade attacks with 1.12% false positives
  - Decoy files can be used to detect all masqueraders within 10 minutes

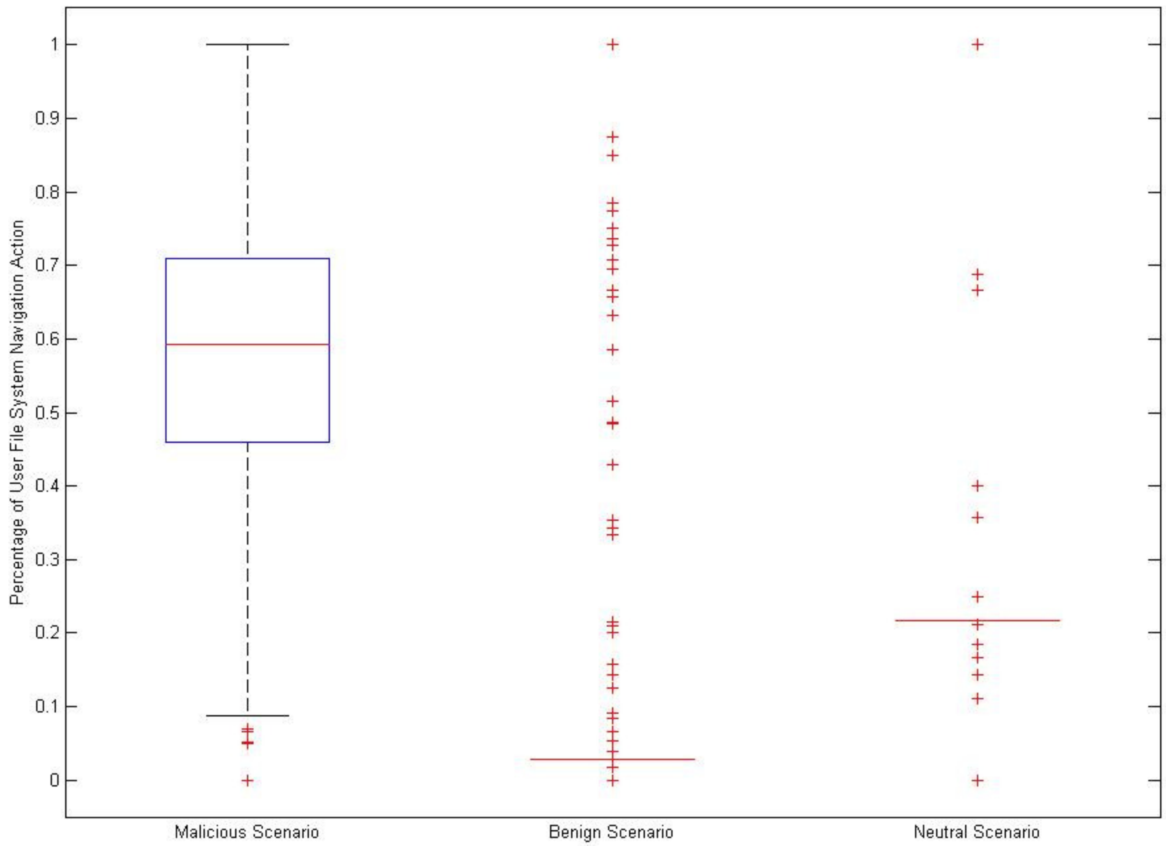
RUU SAMPLE RECORD: REGISTRY ACCESS

Column[1]	Value
Syshash	0cc7ebd580b39bb037627c2a71c979
Auditaction	QueryValue
Processname	explorer.exe
Path	HKCR\CLSID\871C5380-42A0-1069-A2EA-08002B30309D\ShellFolder\Attributes
Stringreturn	SUCCESS
PID	408
PPID	-1
Timestamp	2009-12-09 21:05:46

RESULTS

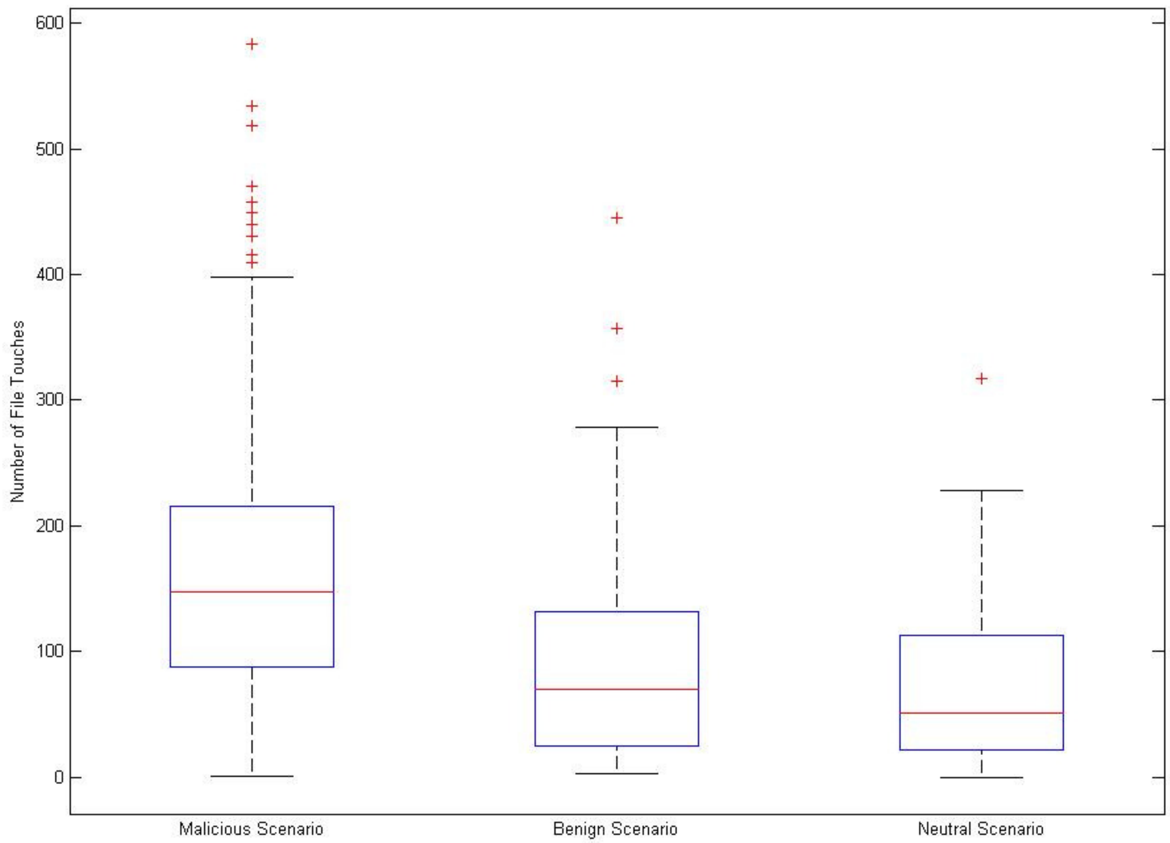
- Search behavior can be used to reveals attacker's malicious intent
- User search behavior profiling achieves100% detection rate of masquerade attacks with 1.12% false positives

RESULTS: DISTRIBUTION OF FILE SYSTEM NAVIGATIONS ACTIONS



BEN-SALEM, M., AND STOLFO, S. J. Modeling user search-behavior for masquerade detection. In To Appear in the Proceedings of the 14th International Symposium on Recent Advances in Intrusion Detection (Heidelberg, September 2011), Springer.

RESULTS: DISTRIBUTION OF FILE TOUCHES

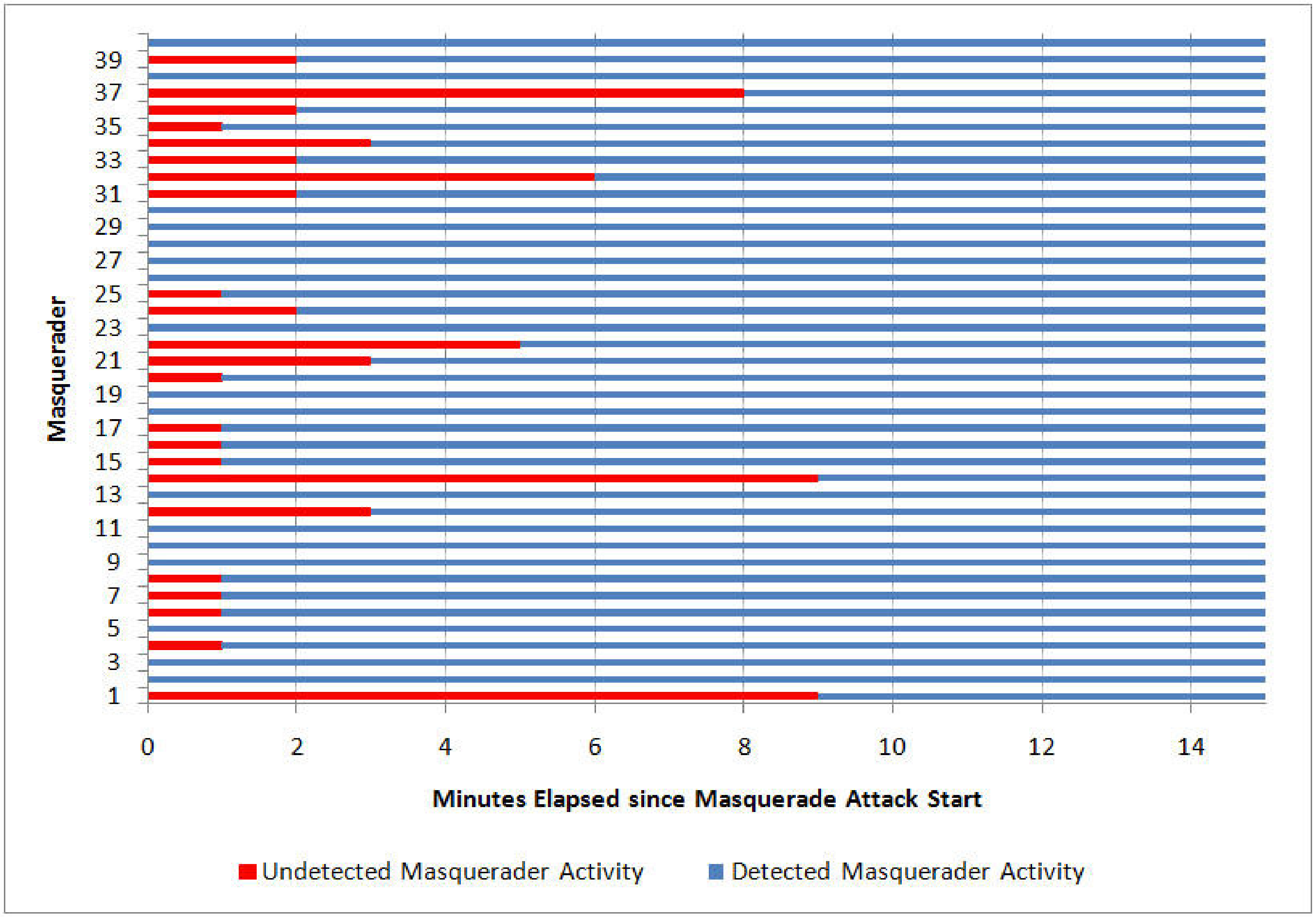


BEN-SALEM, M., AND STOLFO, S. J. Modeling user search-behavior for masquerade detection. In To Appear in the Proceedings of the 14th International Symposium on Recent Advances in Intrusion Detection (Heidelberg, September 2011), Springer.

RESULTS

- Search behavior can be used to reveals attacker's malicious intent
- User search behavior profiling achieves100% detection rate of masquerade attacks with 1.12% false positives
- Decoy files can be used to detect all masqueraders within 10 minutes at most
- More than 40% of masqueraders detected during the first minute of their fraudulent activity

RESULTS: DECOY ACCESS MONITORING

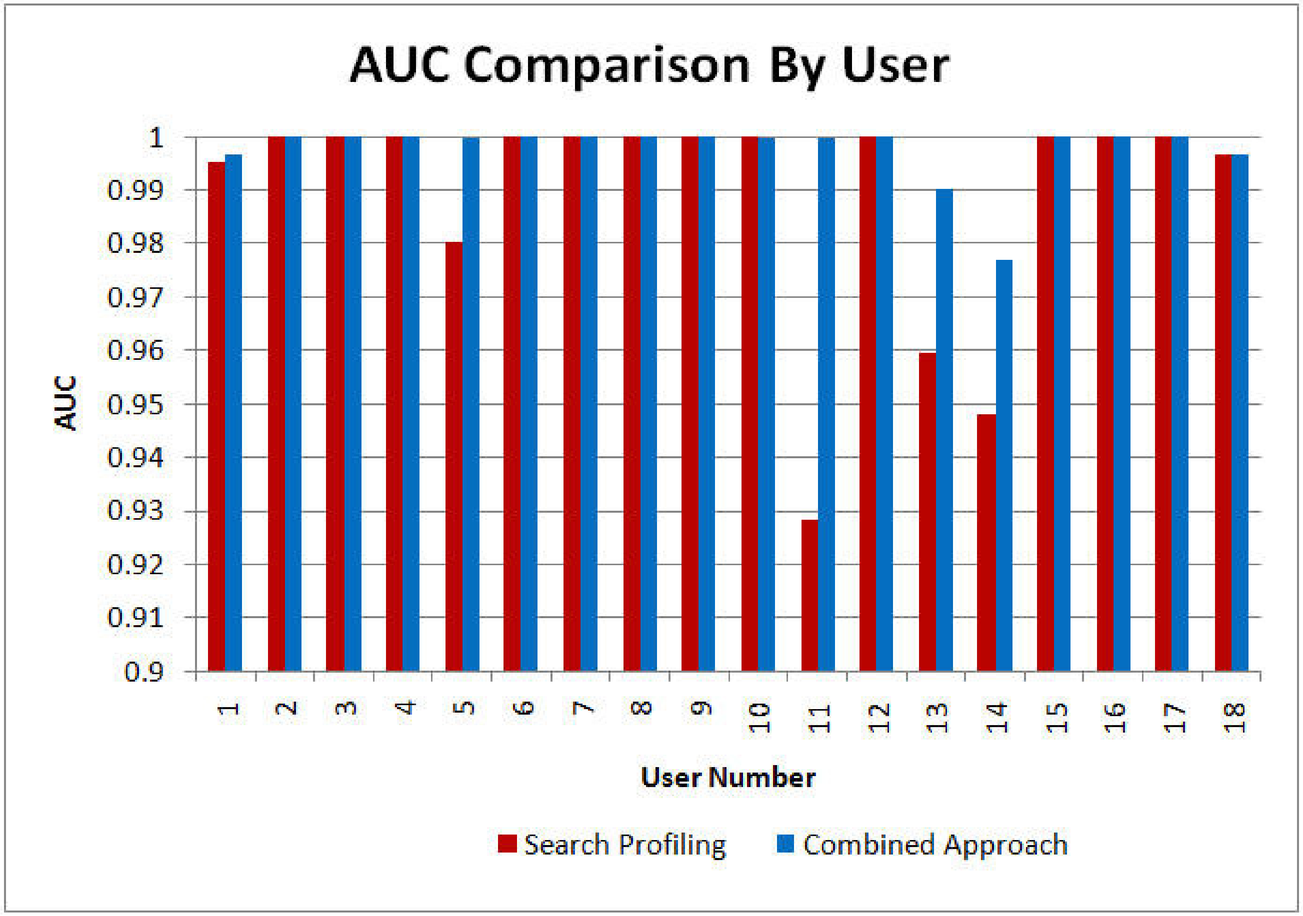


BEN-SALEM, M., AND STOLFO, S. J. Decoy document deployment for effective masquerade attack detection. In DIMVA'11: Proceedings of the Eighth Conference on Detection of Intrusions and Malware & Vulnerability Assessment (Heidelberg, July 2011), Springer, pp. 35 – 54.

RESULTS

- Search behavior can be used to reveals attacker’s malicious intent
- User search behavior profiling achieves100% detection rate of masquerade attacks with 1.12% false positives
- Decoy files can be used to detect all masqueraders within 10 minutes at most
- More than 40% of masqueraders detected during the first minute of their fraudulent activity
- Combining decoys monitoring with search behavior profiling improves accuracy when compared to search profiling alone

RESULTS: SEARCH PROFILING & DECOY ACCESS MONITORING



BEN-SALEM, M. AND STOLFO, S. J. Combining a baiting and a user search profiling techniques for masquerade detection. In Columbia University Computer Science Department, Technical Report # cucs-018-11 (2011).

LESSONS LEARNED

- Compliance-related
  - Initiate IRB review early
  - List a larger sample of user study subjects
  - Have users sign waivers
- Scientific
  - List all assumptions made about users in study scenarios
  - Think carefully about ways for reducing variability and baselining users
  - Perform a power analysis
- Practical
  - Anticipate technology market trends
  - Pilot experiment
  - Have participants fill post-experiment questionnaires

REFERENCES

KEPPEL, G. Design and analysis : a researcher’s handbook. Pearson Prentice Hall, 2004.

MILGRAM, S. Obedience to Authority: An Experimental View. Harpercollins, New York, January 1974.

PEARSON, E. S., AND HARTLEY, H. O. Charts of the power function for analysis of variance tests, derived from the non-central F-distribution. Biometrika 38, 1 (July 1951), 112–130.