COMPUTER SECURITY CLINICAL TRIALS
Lessons learned from a 4-month pilot study

Fanny Lalonde Lévesque, École Polytechnique de Montréal
José M. Fernandez, École Polytechnique de Montréal

CSET’14 - July 2014
1 / Introduction
2 / Computer security clinical trials
3 / Lessons learned and Recommendations
4 / Conclusion and Future Work
OUTLINE

1 / Introduction
2 / Computer security clinical trials
3 / Lessons learned and Recommendations
4 / Conclusion and Future Work
INTRODUCTION

- Security products are an important line of defence against current threats
- The success of malware depends upon both technical and human factors
- Users are involved in both the infection and protection process
- Security products should be tested:
  - With actual users that interact with the product
  - In a realistic environment
  - Over an extended period of time
• Security solutions would be installed and monitored on systems in regular use by regular users

• Data would be gathered on:
  • The performance of the security product in protecting the system
  • How the user interact with the system during this time period

• We performed a 4-month pilot study that involved 50 subjects in order to:
  • Develop and test the validity and viability of the methodology
  • Determine how malware infects computer systems and characterize sources of malware infections
  • Determine how factors such as the configuration of the system, the environment in which the system is used, and user behavior affect the probability of infection of a system
• Protocol certification and approval
  • Computer security risks evaluation committee
  • Research ethics review committee

• Computer risk
  • Guarantee that subjects were not going to be significantly more at risk than the average home computer user who had installed an up-to-date reputable antivirus
  • Protect the university’s IT infrastructure as the experiment implied manipulation of malware files

• Ethics and privacy considerations
  • Guarantee the anonymity of the participants
  • Guarantee confidentiality of the data collected
• Equipment:
  • We provided the same laptop to the subjects
  • All laptops had an identical configurations (OS, tools, AV product, etc)

• Data collected with scripts:
  • List of applications installed and applications for which updates are available
  • Number of Web sites visited and type of Web sites visited
  • Number and types of files downloaded
  • The list of different locations from which the laptop establish connected to the Internet
  • The list of plug-ins installed
  • Number of different hosts to which to laptop communicated
  • Number of hours per day the laptop is connected to the Internet
  • Number of hours per day the laptop is on
• Subject recruiting by advertising the experiment on the University campus

• Users were required to attend 5 in-person sessions

   Initial session:
   • Obtain informed consent from users
   • Provide laptop to users
   • Invite users to complete initial questionnaire (gender, age, status, field of activity, etc.)

   Monthly sessions:
   • Invite users to answer on-line questionnaire to:
     • Assess experience and opinion of the AV product
     • Gain insights about how the computer was used
     • Determine users’ level of security awareness and measure of due diligence they exert to secure their computer
   • Collection of statistical data
Monthly sessions:

• Analysis of the computer to find malware missed by the AV product
  • Tools: HijackThis, ProcessExplorer, Sigcheck, Autoruns, SpyBHORemover, SpyDLLRemover, Whatchanged, Winprefectview
  • Elements were classified using external online resources: www.systemlookup.com, www.processlibrary.com, VirusTotal, Anubis, etc.
  • Files identified as suspicious or dangerous were subject to an in-depth investigation

• If a malicious file is identified by the AV product or our protocol:
  • User is asked to answer a questionnaire to collect more information regarding the potential means and sources of the infection, and any behavioural changes observed on the computer
  • Additional consent is requested from user to collect specific data, such as the browser history, the tshark log files, and the suspected file(s)
Final session:
• Similar to the other monthly sessions
• Users had to answer a post-experiment questionnaire to:
  • Assess their overall experience in the study
  • Help us ide
• Users were asked to keep the collected data for an additional period of 3 months in the event we might need to perform more in-depth analysis on the computer
• We provided users procedures to:
  • Stop the automatic collection of the data
  • Delete the data and tools we installed
  • Reinstall the operating system, if they wanted to do so
OUTLINE

1 / Introduction
2 / Computer security clinical trials
3 / Lessons learned and Recommendations
4 / Conclusion and Future Work
Estimate the required population size before the study

- The required population size is based on many factors:
  - Form of the experiment
  - Hypotheses to be tested
  - Desired power of experiment results
  - Budget available

- Power analysis can help estimate the number of subjects based on:
  - The desired power (indicator of how statistically significant the results may be)
  - The desired statistical level of significance
  - The effect to be measured (estimated based on prior studies, literature review, logical assertion and conjecture)
LESSONS LEARNED AND RECOMMANDATIONS

Select users on a scientific and ethical basis

- Specify the target population using appropriate inclusion and exclusion criteria
- Recruit users based on their characteristics depending on the hypotheses to be tested
Identify relevant independent variables

• Choose the data to be collected based on the research questions

• Identify appropriate independent variables that may affect the dependent variable

• How? Produce an Ishikawa diagram (aka fishbone diagram or cause-and-effect diagram) to identify potential factors
Select the appropriate data collection method

• The collection method should depend on:
  • The data you want to collect
  • The research questions you want to address
  • The type of data analysis you want to perform

• Example: We collected the number of applications installed thought surveys and scripts
  • Average number of applications installed from surveys = 19.02
  • Average number of applications installed from real data = 67.78
  • Recommendation: survey results should be cross-checked with another collection source before analysis
Understand the data before the analysis

• Collected data should be subject to an in-depth investigation to detect and understand any anomalies such as:
  • Missing data
  • Outliers (residual analysis)

• Example: One user had more than 100 detections from the AV product when all others had less than 30 detections
  • Cause: An infected file that the AV was not able to remove, which resulted in multiple detections
  • User has been considered as an outlier and was removed for the analysis
LESSONS LEARNED AND RECOMMANDATIONS

Ground truthing and in-depth analysis

• Number of malware encounters: provided by the AV

• Number of malware infections: based on our protocol
  • Could have been underestimated

• Type of malware
  • Malware classification is inconsistent from one AV vendor to another
  • The problem is even worse with unwanted software (adware, spyware, etc.)
  • Solution: in-depth forensics analysis based on the collected data
Evaluate potential bias before the study

- Controlling factors can reduce experimental bias, but they can also reduce the ecological validity of the results

- We had to control system configuration variables that could affect AV performance in protecting the system
  - Users had the same laptop, same OS, same AV, same tools, etc.
  - They not allowed to install another AV, uninstall or disactivate the AV being evaluated

- We decided to evaluate the bias related to user behavior and environment, rather than control it
  - Users share their computer, use VM, use a private navigation mode, etc.
  - We asked the users at the end of the experiment how they may have changed their behavior knowing that they were participating in the study
OUTLINE

1 / Introduction
2 / Computer security clinical trials
3 / Lessons learned and Recommendations
4 / Conclusion and Future Work
• We showed that computer security clinical trials are a viable and complementary alternative to traditional testing

• Designing and conducting such studies presents many scientific, technical and ethical challenges

• Computer security clinical trials could:
  • Help AV vendors to improve their products and understand how users adopt security solutions and react the security threats
  • Be applied to evaluate the effectiveness of training and education on users
  • Help key deciders in IT to make better evidence-based decisions
Results were presented in:

- « A clinical study of risk factors related to malware infections » , ACM CCS 2013

Contact:

José M. Fernandez  
Associate professor at Polytechnique Montréal  
Department of Computer Engineering  
Email: jose.fernandez@polymtl.ca