Report for 06-How to Design Computer Security Experiments

Xinyu Chen

March 2, 2017

1 Summary

This paper discussed scientific principles for designing computer security experiments. The authors argue that it is important to apply scientific methods to computer science experiments. The most relevant qualities regarding the experiment designs are *Falsifiable*, *Controlled* and *Reproducible*. Another important quality they mentioned is *Blind* and *Double Blind* so that the experimenters and subjects are not aware that certain subjects are controls or not. This quality can prevent the bias introduced by experimenters or subjects. In the following sections, the paper explained the first three qualities in detail.

The quality of *Falsifiable* should have *observability* and *measurability*. To achieve this, a hypothesis needs to give specific definitions such that they can be observed and measured. The quality of *Controlled* distinguish observations and experiments. Researchers need to set up tests where they can isolate the interesting independent and respond variables. Only by this, they can interpret and establish causal relationship from their experiments. The quality of *Reproducible* is to document conditions and save dataset, so other researchers can reproduce and validate the experiments. An example of testing the performance of a new firewall illustrated the authors' arguments.

2 Key Takeaway

The three qualities of *Falsifiable*, *Controlled* and *Reproducible* are the essence of this paper. They also gave some scenarios to illustrate good and bad experiment designs. This are useful for our future researches, not limited to computer security problems but also broader research areas.

3 Discussions

The paper is straight forward and general, there is not much questions.

- *Input Data*. The authors emphasized to validate input data. They also mentioned the best way to characterize their data set is to release it. Let others be able to use it and reproduce their experiment. This seems to be a subjective method.

- *Blind and Double Blind*. The authors mentioned the quality of *blind* and *double blind* are important when humans are subjects. In computer security and a lot of computer science areas, humans are important factors. What is an example that require *blind* or *double blind* in computer science experiments?

- *Machine learning Scenarios*. What are some examples that can illustrate the three principles of *Falsifiable*, *Controlled* and *Reproducible* in machine learning experiments? It seems machine learning algorithms already have apply confusion matrix, recall, precision, accuracy rate and visualization techniques to apply the above principles.