Review 22

4/12/12

CS6V81.502

Seungtack Baek

**Automatic Input Rectification**

In the second half of this Paper, Long et al. actually evaluated the tool they developed for input rectification. They evaluated SOAP (the name of the tool) by testing SOAP on 5 different applications. They had 6 malicious inputs and multiple benign inputs collected by web crawler.

As a result, SOAP was able to detect all malicious inputs, rectify and successfully passed those into the application. Moreover, they did not find any vulnerability exploit from these rectified inputs (So, in a way, the malicious inputs become sanitized).

In this evaluation they tried to show the broad application of SOAP by selecting subject applications that takes different file types as their input, such as image, video and audio files. (One thing I liked to see was .pdf files, but oh well..) All of the malicious files of these types could impose possible overflow and SOAP was able to fix this issue.

Not only SOAP nullified the malicious input, it also tried to preserve as much information as possible and they were able to preserve most of the data. PNG and WAV were preserved 100% and only about 0.5% of the data were lost for JPEG and TIFF.

After this, they also empirically found the optimal size of training set. They started from some arbitrary size and then increased by 200 per each run. They found that the lager the training set is, the lees data they lost (which is as they expected).

Yet, the figure 7 and 8, which shows the result for the evaluation, were little bit hard to see. They tried to show too much in a table. I wish that they split the table into some correlated groups.

**Question**

1. What is mechanical-turk evaluation?
2. What is web crawler?