NCSU Science of Security Lablet
Transitional Relevance and Research Summary Statement

Name: Hui Guo

Hard Problem Area: Extracting norms from textual requirements and regulations

Overall Research Goal(s):
Ultimately I hope to show with my research that normative relationships can be extracted from requirement and regulation documents efficiently, which in turn can be used to build normative models automatically. My goal is to build and test a systematic process of norm extraction from text. I also want to investigate how to refine and utilize the extracted norms in order to speed up the transformation from textual documents to normative models.

Broad Impact:
Norms (commitments, authorizations and prohibitions) can help express requirements precisely [1] and characterize potential breaches by identifying violation conditions [2]. Normative models based on manually extracted norms have been used for capturing variability of law and checking regulatory compliance [3]. Previous work on norm extraction was primitive and the results might not be used directly for model building [4]. I want to make the norm extraction process more efficiently and effectively.

Specific Research Goals:
1. Investigate and identify the peculiarity of norm extraction from requirements and regulations.
2. Improve and extend previous work on norm extraction. To start with, study the efficiency and effectiveness of crowdsourcing on norm extraction.
3. Find ways to refine extracted norms with the intend to use them in normative model building.

Proposed Data Collection (if applicable):
I will need textual requirement and regulation documents as input. I will need to collect extracted norms (e.g. types and elements) as training sets.

Success Criteria:
We can determine the success of this research by examining the extracted norms and checking the precision and recall of this retrieval. For a proper normative model building process, the extracted norms should be able to be used as input.

Anticipated Difficulties, Limitations, and Criticisms:
1. Norm extraction from text could be very domain-specific. I plan to start with a carefully selected data set.
2. The accuracy of the process may rely on a large training set. Crowdsourcing the labeling task could be a solution.
3. I’m not sure how extracted norms can be refined to build models. I will need to investigate the challenges.
References:


