Research Plan Review:

SECURITY NORM EXTRACTION FROM TEXTUAL REQUIREMENTS AND REGULATIONS

Hui Guo

Faculty: Dr. Munindar Singh

01/24/2017
Outline

• Justification
• Previous Work
• Research Plan
Justification

• Norms and normative models
  • can be used to formalize security requirements and breaches;
  • should be generated from text automatically.

• Norm extraction from text is nontrivial.
Previous Work

• Norm extraction
  • From business contracts\cite{1}
  • From regulatory documents\cite{2}

• Normative model building
  • Nòmos 2\cite{3}

• Logic Modeling for requirements
  • Production rule models\cite{4}\cite{5}

\begin{footnotesize}
\begin{enumerate}
\item Xibin Gao and Munindar P. Singh, \textit{Extracting Normative Relationships from Business Contracts}, AAMAS 2014
\item Mustafa Hashmi, \textit{A Methodology for Extracting Legal Norms from Regulatory Documents}, EDOCW 2015 IEEE
\item Jeremy C. Maxwell and Annie I. Antón, \textit{Checking Existing Requirements for Compliance with Law Using a Production Rule Model}, RELAW 2009
\item Jeremy C. Maxwell and Annie I. Antón, Developing Production Rule Models to Aid in Acquiring Requirements from Legal Texts, International Requirements Engineering Conference 2009 IEEE
\end{enumerate}
\end{footnotesize}
Research Plan

• Investigate norms in requirements
• Norm extraction
  • Input: requirement/regulation sentences
  • Identify:
    • Does it contain a norm?
    • If yes, what type is it?
  • Extract:
    • Subject (holder), object (beneficiary)
    • Antecedent, consequent
  • Refine results for normative model building
• Crowdsourcing
Examples

Sentence: TieTech warrants and represents that it shall strictly adhere to the Product Specifications set forth in Appendix III attached hereto and by reference made a part hereof[1].
- Norm type: dialectical commitment
- Subject: TieTech
- Object: null
- Antecedent: true
- Consequent: TieTech warrants and represents that it shall strictly adhere to the Product Specifications set forth in Appendix III attached hereto and by reference made a part hereof

Sentence: To avoid a loss of visibility and protection against intra-host attacks, duplication of the physical network protection capabilities may be required on the virtual network[2].

Sentence: Organizations should consider the risk and performance tradeoffs between having traffic hidden within the hypervisor versus exposing that traffic to the physical network for monitoring[2].

[2] NIST Special Publication 800-144
Norm Extraction: Crowdsourcing

• Justification
  • Norm extraction is an NLP/information retrieval task;
  • Crowdsourcing is reliable and efficient\textsuperscript{[1]}.

• Previous work
  • NER using crowdsourcing\textsuperscript{[2]}

\textsuperscript{[2]} Mark Dredze, \textit{Annotating Named Entities in Twitter Data with Crowdsourcing}, CSLDAMT 2010
Norm Extraction: Crowdsourcing (cont.)

• Research plan\textsuperscript{[1]}:
  • Evaluation
    • Which processes can be crowdsourced?
  • Task design
    • How can the tasks be decomposed into microtasks?
  • Integration
    • How can work be integrated with existing processes?

\textsuperscript{[1]} Olsen Tim and Carmel Erran, \textit{The Process of Atomization of Business Tasks for Crowdsourcing}, Strategic Outsourcing: An International Journal 2013
Norm Extraction: Crowdsourcing (cont.)

• **Thoughts**
  • How to control quality?
    • Existing methodology as a baseline
  • How can the results be used later?
    • As a training set?
    • As inputs for logic modeling?
  • Time constraints
Normative Modeling

• How can the extracted norms be refined?
  • Pronominal co-reference
  • Atomization of antecedents and consequents
  • Identification or co-reference of antecedents and consequents

• Dilemma
  • Size of a training set (the bigger the better)
  • Size of a normative model (the smaller the better)
Thank you!

Happy New Year!
(The Year of Chicken starts on Jan 28th, 2017)