**Problem Statement**

- Access control policies (ACPs) should be expressed correctly because improper policy expression introduces security vulnerabilities.
- ACPs are embedded in the security requirements document in natural language.
- ACPs should be derived from security requirements and converted to machine-executable instructions.
- Manual extraction is tedious, complex, expensive, labor-intensive, and error-prone.
- Research Question: How do we automatically extract NGAC Policy from security requirements documents?

**Research Question:** Manual extraction is tedious, complex, expensive, labor-intensive, and error-prone.

**Phase 1: ACPs Identification**

- Classify whether a sentence is a policy or not.
- Binary classification

**Phase 2: Semantic attribute mapping**

- Identifying the semantic roles of the words
- Mapping semantic roles with NGAC attributes

**The Proposed Approach**

**Figure 1:** The proposed Approach Flowchart

**NGAC Association**

- (\(UA \rightarrow \{\text{ars}\} \rightarrow OA\))
- NGAC Association in terms of SRL Tags:
  - (Arg0 \(\rightarrow V \rightarrow Arg1\))
  - Where Arg0 is UA, \(V\) is ars, and Arg1 is OA

**NGAC Association Example:**

- Input: The patient can add the HCP to their list of providers.
- Output: (patient, add, HCP)

**NGAC Prohibition**

- \(\text{deny}(UA \rightarrow \{\text{ars}\} \rightarrow OA)\)
- NGAC Prohibition in terms of SRL Tags:
  - \(\text{deny}(Arg0 \rightarrow V \rightarrow Arg1)\)
  - Where Arg0 is UA, \(V\) is ars, and Arg1 is OA

**NGAC Prohibition Example:**

- Input: The LHCP is not able to edit any past appointments.
- Output: deny(LHCP, edit, past appointments)

**NGAC Obligation**

- (\(\text{Event}(UA \rightarrow \{\text{ars}\} \rightarrow OA)\) → \(\text{Response}(op \rightarrow P)\))
- NGAC Obligation in terms of SRL Tags:
  - \((\text{ARG } MP, \text{ ARG } ADV )\) → \((\text{TMP} \rightarrow \text{ADV}(Arg0 \rightarrow V \rightarrow Arg1)\)
  - Where \(\text{TMP}\) is the event, and \(\text{ADV}\) is the response (Arg0 is UA, \(V\) is ars, and Arg1 is OA)

**NGAC Obligation Example:**

- Input: If a patient has not taken an office visit satisfaction survey, the patient may take the survey for an office visit.
- Output: \(\text{TMP}\) (If a patient has not taken an office visit satisfaction survey, \(\text{ADV}\) patient may take the survey)

**Datasets**

- **iTrust-v1** (DS1): healthcare dataset
- **iTrust-v2** (DS2): largest version of DS1
- **IBM (DS3)**: IBM course registration system
- **CyberChair (DS4)**: conference domain
- **Collected-ACPs (DS5)**: papers, websites

**Table 1:** Size of Training Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>ACP</th>
<th>Not ACP</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>419</td>
<td>53</td>
</tr>
<tr>
<td>DS2</td>
<td>552</td>
<td>609</td>
</tr>
<tr>
<td>DS3</td>
<td>179</td>
<td>223</td>
</tr>
<tr>
<td>DS4</td>
<td>141</td>
<td>163</td>
</tr>
<tr>
<td>DS5</td>
<td>141</td>
<td>163</td>
</tr>
</tbody>
</table>

**Table 2:** Accuracy Report of FastText Classification

- **Classification:** as ACP or not ACP
- **FastText:** open-source for learning word embeddings and word classifications

**Semantic Role Labeling (SRL):**

- Determines the semantic relations between a predicate and its associated participant
- Tags are used (Arg0, Arg1, Arg2, TMP, LOC, DIR, MNR)
- Mapping SRL tags with the NGAC attributes in ACP sentences semantically as (user attribute (UA), access rights (ars), and object attribute (OA))

**AllenNLP:**

- Adapting a transition-based neural network
- Achieve state-of-the-art performance

**Contributions**

- An approach to identify ACP sentences in natural language documents
- An automated approach to extract the NGAC attributes in each ACP sentence using SRL

**References**


**Appendix B:**

- FastText source for healthcare dataset
- DS3: IBM course registration system
- DS4: Cyberspace – Cybersecurity and Cybercrime Conference domain
- DS5: Collected-ACPs papers, websites

**Table 3:** Size of Training Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>ACP</th>
<th>Not ACP</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>419</td>
<td>53</td>
</tr>
<tr>
<td>DS2</td>
<td>552</td>
<td>609</td>
</tr>
<tr>
<td>DS3</td>
<td>179</td>
<td>223</td>
</tr>
<tr>
<td>DS4</td>
<td>141</td>
<td>163</td>
</tr>
<tr>
<td>DS5</td>
<td>141</td>
<td>163</td>
</tr>
</tbody>
</table>

**Table 4:** Accuracy Report of FastText Classification

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>0.88</td>
<td>0.96</td>
<td>0.92</td>
</tr>
<tr>
<td>DS2</td>
<td>0.94</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>DS3</td>
<td>0.66</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td>DS4</td>
<td>0.91</td>
<td>0.96</td>
<td>0.93</td>
</tr>
<tr>
<td>DS5</td>
<td>0.92</td>
<td>0.97</td>
<td>0.94</td>
</tr>
</tbody>
</table>