Automated Construction of NGAC Policy from Natural Language

Saja Alqurashi, Videep Venkatesha
Advisor: Prof. Indrakshi Ray
Co-Advisor : Hossein Shirazi
Colorado State University Fort Collins, Colorado
San Diego State University, San Diego, CA

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 ?

The Proposed Approach

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

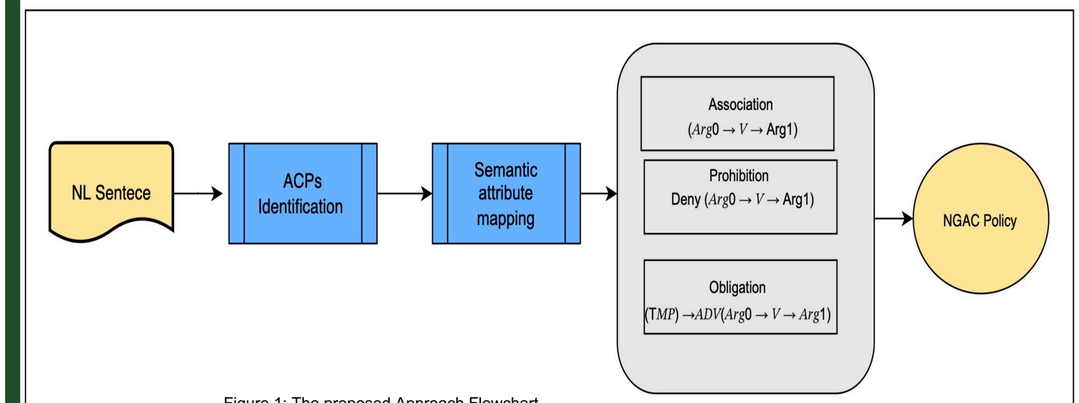

Figure 1: The proposed Approach Flowchart

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

	ACP	Not ACP
DS1	419	53
DS2	552	609
DS3	179	223
DS4	141	163
DS5	141	163

Table 1: Size of Training Datasets

Phase 1: ACPs Identification

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

Datasets	Precision	Recall.	F1-score
DS1	0.88	0.96	0.92
DS2	0.94	0.97	0.96
DS3	0.66	0.73	0.69
DS4	0.91	0.96	0.93
DS5	0.92	0.97	0.94

Table 2: Accuracy Report of FastText Classification

Phase 2:Semantic attribute mapping

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 SRL** :*

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

NGAC Association

- **NGAC Association:**
  $(UA - \{ars\} - OA)$
- 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

- **NGAC Prohibition:**
  deny $(UA - \{ars\} - OA)$
- NGAC Prohibition in terms of SRL Tags:
  $deny\ (Arg0 \rightarrow V \rightarrow Arg1)$

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

- **NGAC Obligation:**
  $(Event:(U\ A - \{ars\} - OA) \rightarrow Response:(op - P)$
- NGAC Obligation in terms of SRL Tags:
  - $(ARG\ T\ MP, ARG\ ADV) : When\ (TMP) \rightarrow ADV\ (Arg0 \rightarrow V \rightarrow Arg1)$

Where *TMP is the event, and 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 for an office visit yet. The patient may take the survey for an office visit.
- **Output**: *TMP* (If a patient has not taken an office visit satisfaction survey), *ADV*(patient may take the survey)

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

[1] Khalid Zaman Bijon, Ram Krishnan, and Ravi Sandhu. 2013. Towards an attribute based constraints specification language. In 2013 International Conference on Social Computing. IEEE, 108–113.
[2] Matt Gardner, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson Liu, Matthew Peters, Michael Schmitz, and Luke Zettlemoyer. 2018. Allennlp: A deep semantic natural language processing platform. arXiv preprint arXiv:1803.07640 (2018), 101611.
[3] Andrew Meneely, Ben Smith, and Laurie Williams. 2012. Appendix B: iTrust electronic health care system case study. Software and Systems Traceability (Feb2012), 425.
[4] Pierangela Samarati and Sabrina Capitani de Vimercati. 2000. Access control: policies, models, and mechanisms. In FOSAD. 137–196.
[5] John Slankas, Xusheng Xiao, Laurie Williams, and Tao Xie. 2014. Relation extraction for inferring access control rules from natural language artifacts. In ACSAC. 366–375.
[6] Richard Van De Stadt. 2012. Cyberchair: A web-based groupware application to facilitate the paper reviewing process. arXiv preprint arXiv:1206.1833 (June 2012).
[7] Yutang Xia, Shengfang Zhai, Qinting Wang, Huiting Hou, Zhonghai Wu, and Qingni Shen. 2022. Automated Extraction of ABAC Policies from Natural-Language Documents in Healthcare Systems. In 2022 IEEE(BIBM). IEEE, IEEE,1289–1296.