Generating Privacy and Security Threat Summary for Internet of Things

Domonique Cox
Dept. of Computer Science Engineering
Prairie View A&M University
dcox15@student.pvamu.edu

Kaiqiang Song
Dept. of Computer Science
University of Central Florida
kqsong@knights.ucf.edu

Dr. Fei Liu
Dept. of Computer Science
University of Central Florida
feiliu@cs.ucf.edu

Abstract

- There is an urgent need to understand website privacy policies. They are sometimes long, vague and hard to understand the information.
- To help with this problem, summarization is an effective technology tool to help customers quickly glance through the information.
- Text summarization is the art of condensing lengthy text documents into a short easy to read article using software.
- Different types of techniques includes extractive and abstractive summarization. My future plan on how summarization can help privacy policies to be more understandable is by annotating data.

Extractive Summarization

- Extractive summarization is to directly select tokens from the original text. It appears in the same order as the original document.

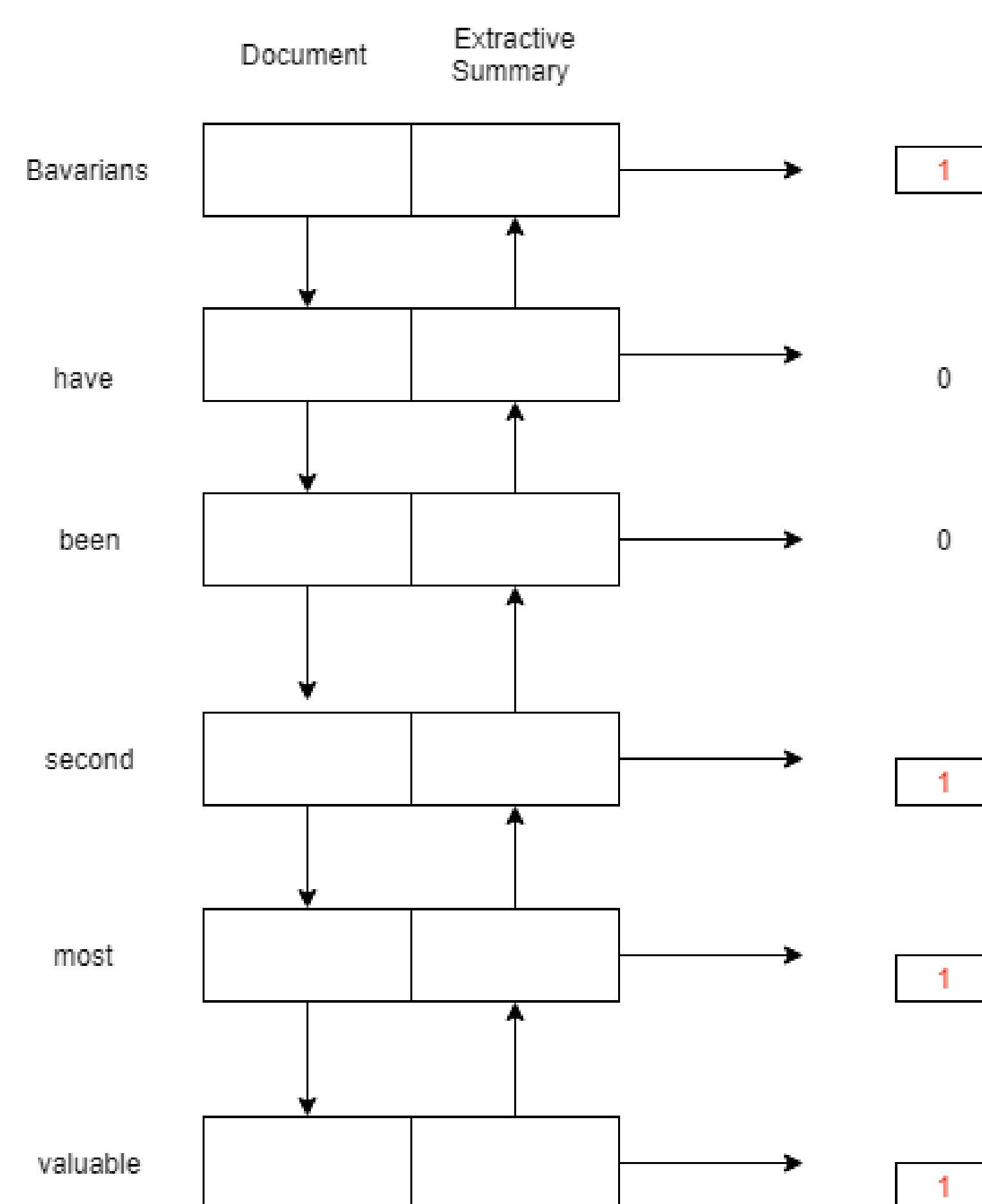| Original Text	**Germany** emerges victorious in 2-0 **win Saturday** against Argentina.
Summary	Germany win Saturday.

Figure 1a : Extractive Summarization

Figure 1b: System framework for an Extractive Summary

Abstractive Summarization

- Abstractive summarization is expressed by the words of the summary author. Rewording phrases and clauses from the original text.

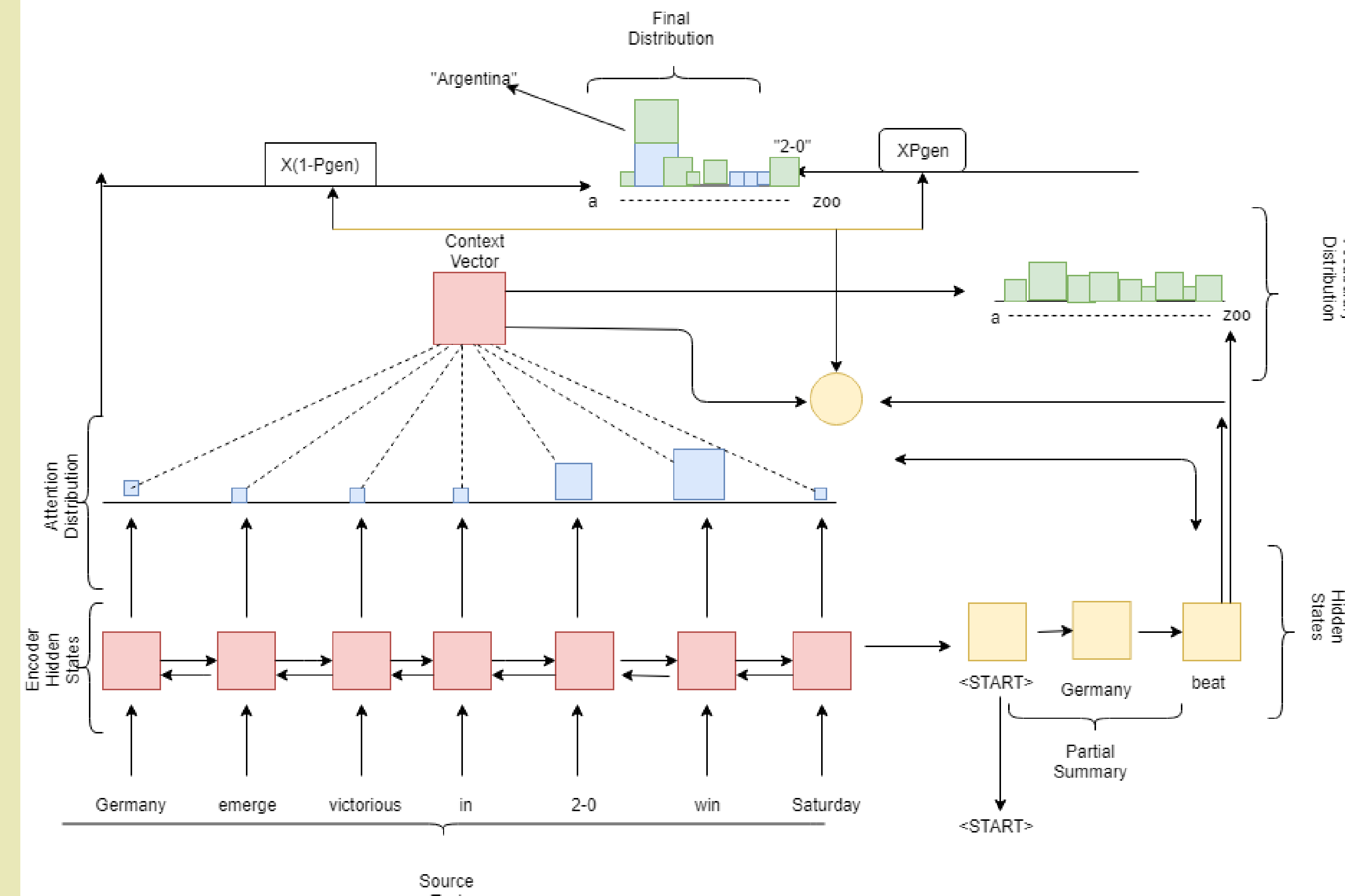| Original Text	Germany emerge victorious in 2-0 win Saturday.
Summary	Germany **beat Argentina**.

Figure 2a: Abstractive Summarization

Figure 2b: Abstractive summary using Pointer Generator Model

References

1. See, Abigail, Liu,Peter and Manning, Christopher."*Get To The Point: Summarization with Pointer Generator Networks*",2007,arXiv:1704.04451[cs.Cl].
2. Arumae, Kristjan, and Fei Liu. Reinforced Extractive Summarization with Question-Focused Reward.2018.
3. Joel R. Reidenberg, Travis Breaux, Lorrie Faith Carnor, and Brian French, Disagreeable Privacy Policies: Mismatches Between Meaning and Users' Understanding , 30 Berkeley Tech. L.J.39(2015).
4. A.Nenkova and K. McKeown, Auto-matic Summarization, Foundation and Trens in Information Retrieval, vol 5, nos 2-3, pp 103-233, 2011.

Future Plan

- To help summarize privacy policies or any summary article to be more understandable is by annotating data.
- Generating short text summaries has been a challenging task for Natural Language Processing researchers.
- One of the solutions is to create a list of questions for each topic of discussion.
- Using question to answer pairs to help validate the informativeness and the correctness of summaries.
- When it becomes time to test, it will be evaluated by how many questions can be answered by the system summary.

Figure 3: Topic and Answers for Annotating Data

Acknowledgments