



Interactive Exploration of Asynchronous Conversations: Applying a User-centered Approach to Design a Visual Text Analytic System

Enamul Hoque, Giuseppe Carenini

{enamul,carenini}@cs.ubc.ca
University of British Columbia

Shafiq Joty

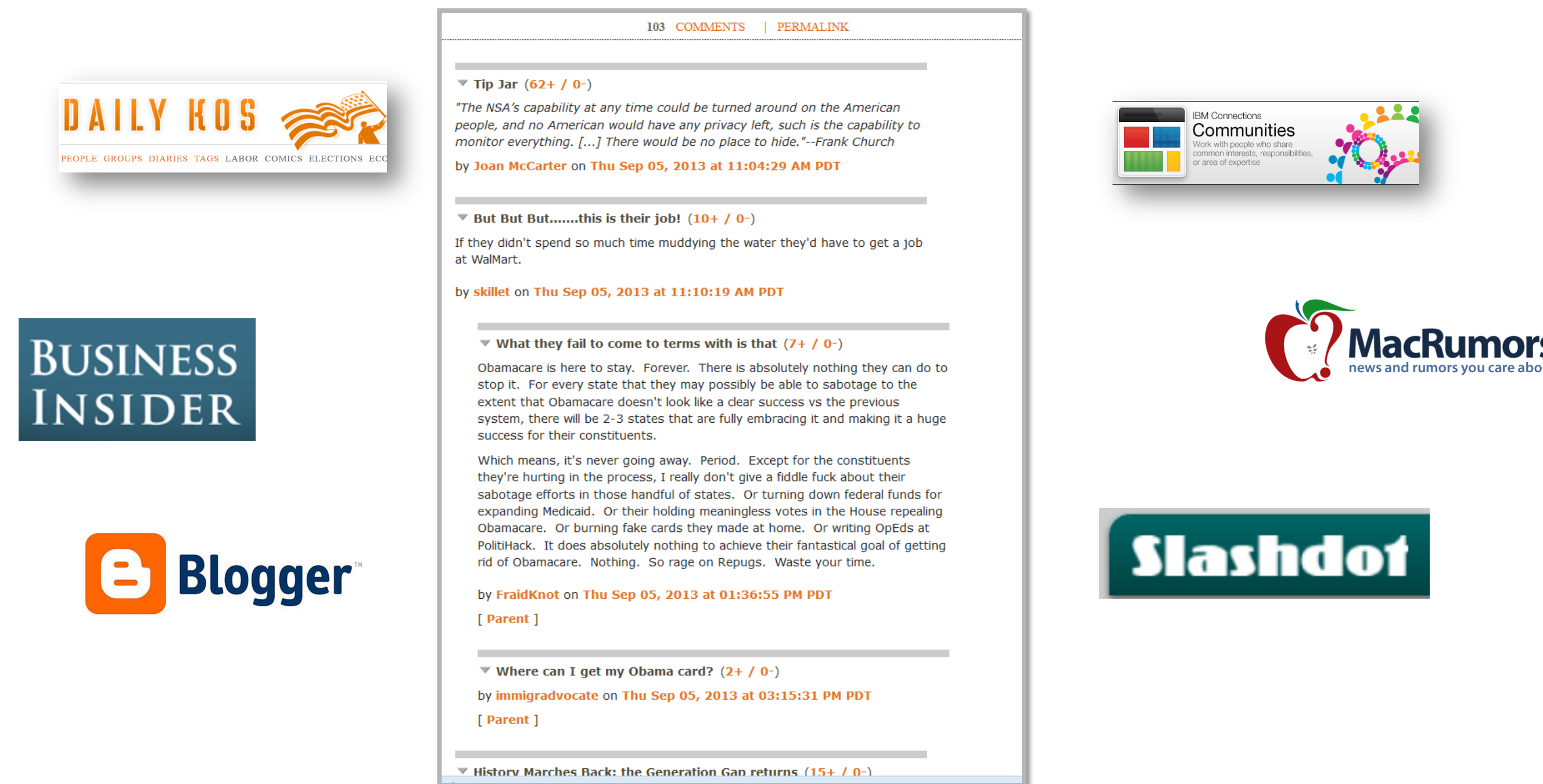
sjoty@qf.org.qa
Qatar Computing Research Institute

The Problem

Asynchronous conversations such as a blog can generate a **long and complex thread** as comments are added by the participants

Information overload:

- ☐ The readers skip comments
- ☐ Generate short response
- ☐ Leave the discussion prematurely



Our Approach

User centered design

Apply Nested Model [Munzner 2009]

- What NLP methods should be applied?
- What metadata are important?
- How the information should be visualized?

Tightly integrate NLP and InfoVis techniques

Applying User Centered Design

Characterizing the Domain of Blogs

Blog Data and Tasks Abstractions

Extracting Data from Blogs

ConVis: Interactive Visualization of Conversations

Why and how people read blogs?

- Variety seeking behaviour
- Skimming behaviour

- Information and guidance seeking
- Keep track of arguments and evidences
- Have fun and enjoyment

Comment

Sample Tasks

Topic

- What this conversation is about?
- What do people say about **topic X**?
- Why are people supporting an opinion?

Thread

Author

Opinion

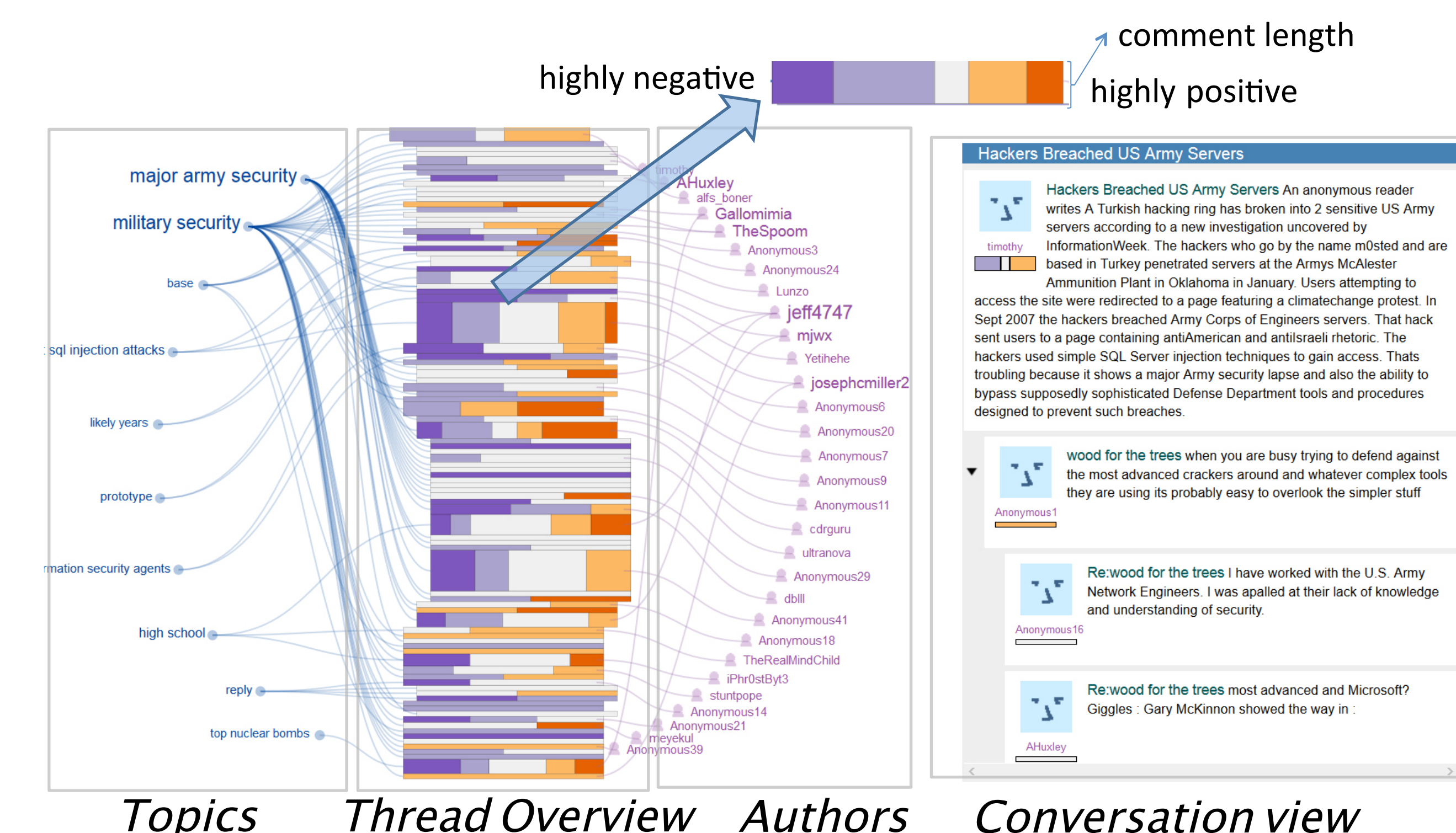
Topic Model

Segmentation: Use graph-based clustering model that considers lexical cohesion and conversational structure.

Labeling: Generate k keyphrases for each segment using graph-based co-ranking method.

Opinion extraction

- Apply Semantic Orientation CALculator
- Compute polarity distribution for each comment



Interactive Topic model to Support User Tasks

Why?

Given an initial model:

- Topic model might be noisy
- Users may be different (e.g., in expertise)
- Task may require to change the topic granularity

Example: A system generated topic is “Military security”.

It consists of sub-topics: “advance hacking”, “defence facility” and “security lapses”.

Task: What opinions are expressed about “**security lapses**”?

The user needs to split “Military security” into further sub-topics.

How?

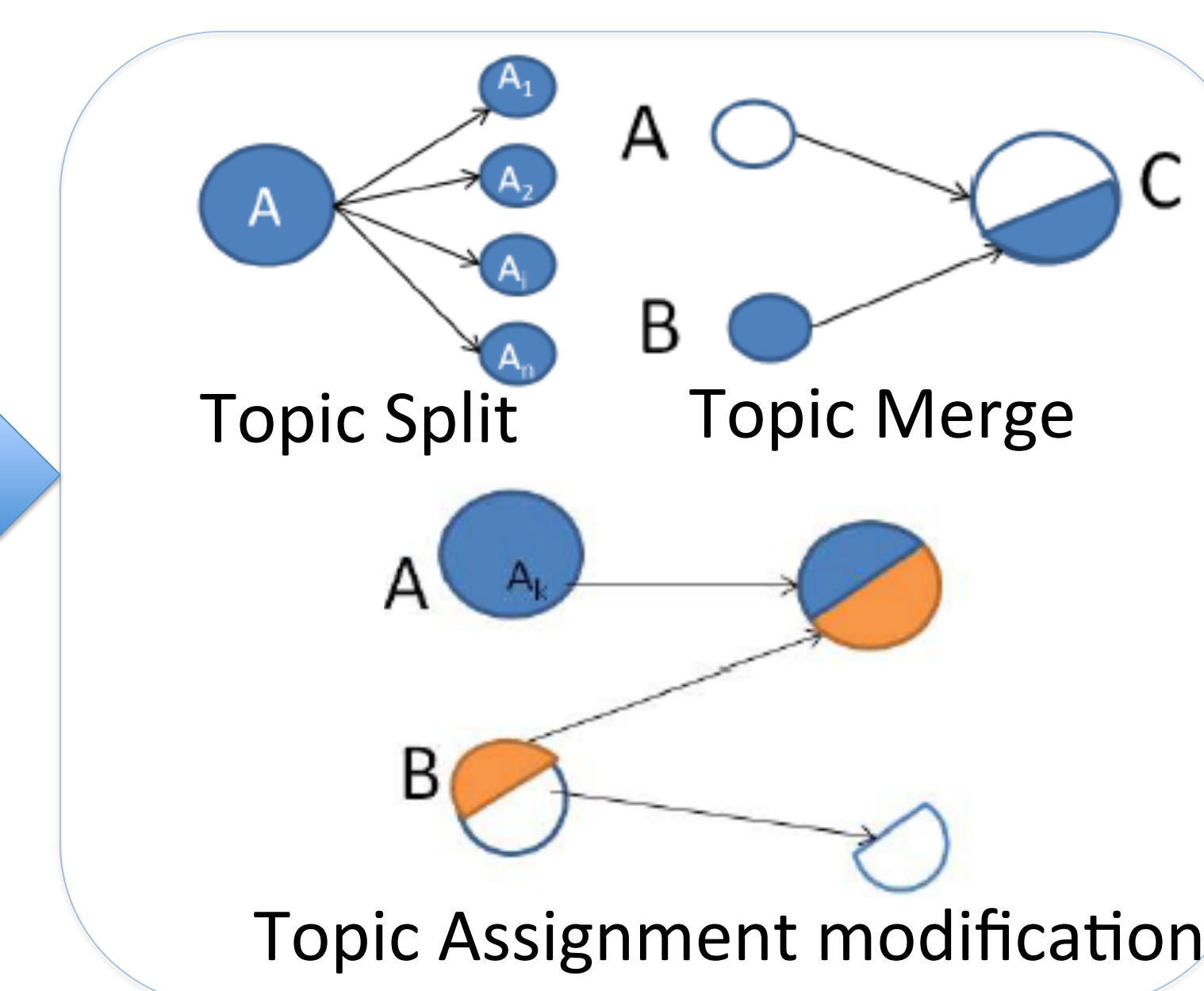
Underlying Topic Model

updates

User feedback

Visual Interface

- Supports user feedback interactions
- Visualizes the changes in the results



Future Directions

- Couple advanced NLP Methods with Interactive Visualizations
- Run a summative evaluation

For live demo and related papers:
<http://www.cs.ubc.ca/~enamul/convis/>

