

# Understanding the Signature of Controversial Wikipedia Articles through Motifs in Editor Revision Networks



James Ashford (Cardiff), Liam Turner (Cardiff), Roger Whitaker (Cardiff), Alun Preece (Cardiff), Diane Felmlee (PSU), Don Towsley (Umass).

## Objectives

- Understand the characteristics of **controversial** content through collaboration.
- Combat misleading narrative in **Wikipedia** articles preventing a neutral consensus from emerging.
- Demonstrate the **effectiveness of motif-based analysis** to reveal **group behaviours**.

## Technical Challenges

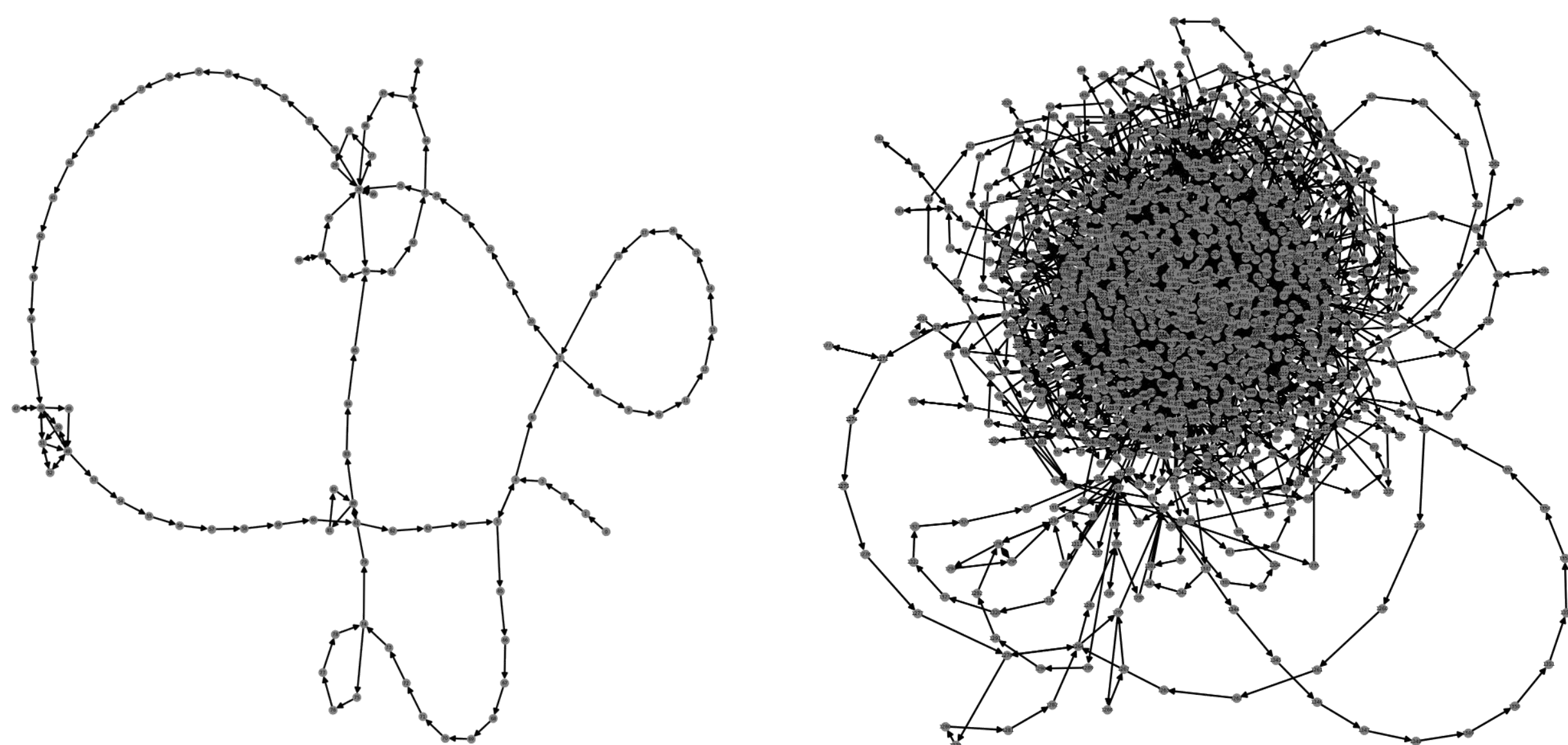
- Processing complexity of emerging content as a cause of cooperation and conflict.
- Identification of controversy without need of complex semantic analysis.
- Representation of article revision logs with many entries.

## Approaches

- Collection of **21,000+ random and controversial articles**. **18,970 non-controversial** articles and **2,661 controversial** articles were processed .
- Generation of **revision networks**. Unique editors (vertices) and directed edges show that an editor added a revision after another editor.
- Use of **motif analysis** to find over and under representations of induced subgraphs using **subgraph ratio profile**.

$$\Delta_i = \frac{N_{real_i} - \langle N_{rand_i} \rangle}{(N_{real_i} + \langle N_{rand_i} \rangle) + \epsilon} \quad SRP_i = \Delta_i / \sqrt{\sum \Delta_i^2}$$

Subgraph ratio profiles are produced by calculating the abundance of each subgraph formation with respect to a random null model.



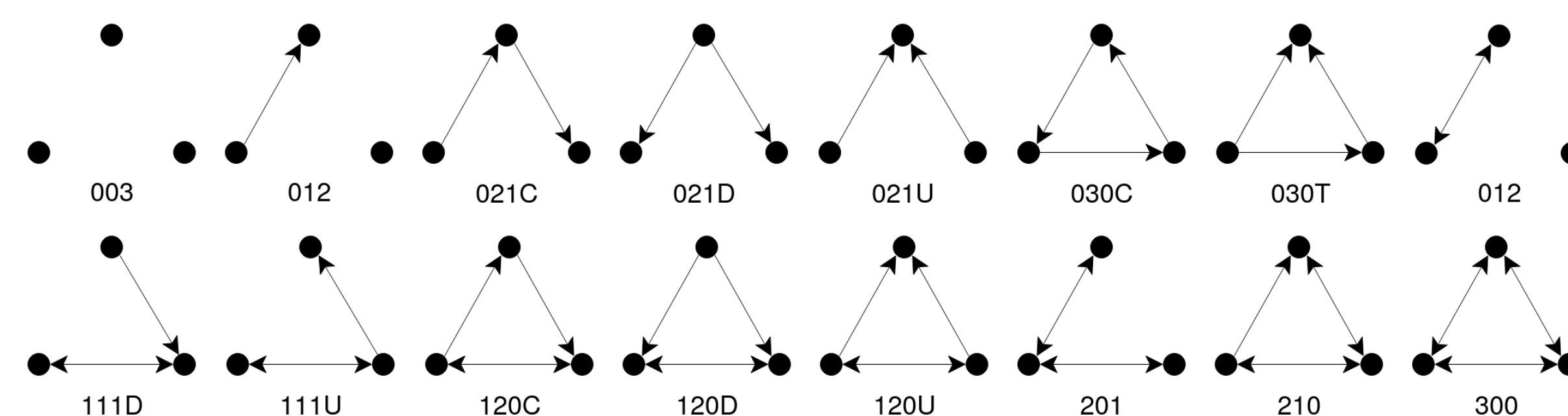
Examples of non-controversial (left) and controversial (right) revision networks.

## Military & Coalition Relevance

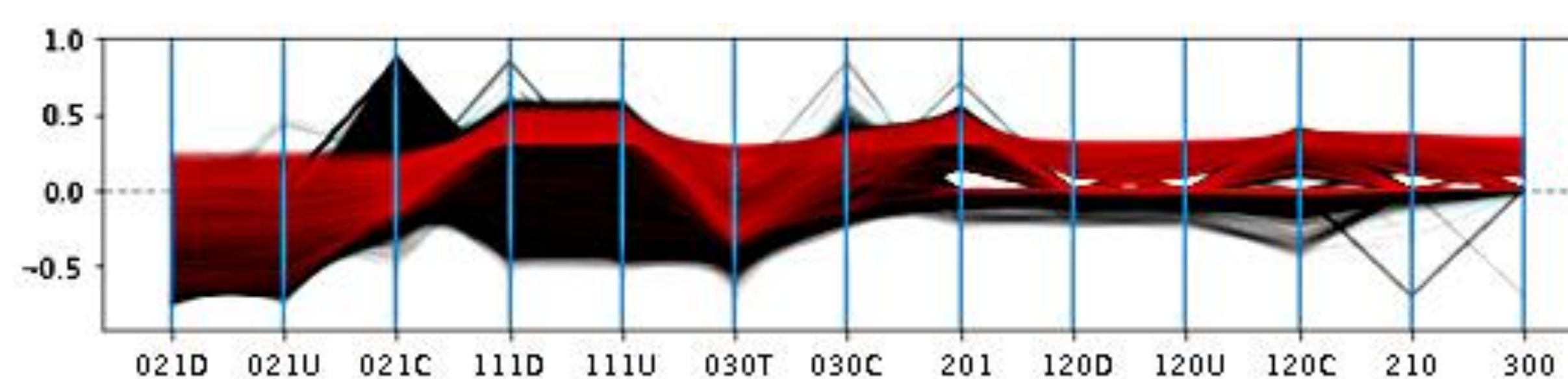
- Understanding the **behaviour-driven** aspects of online collaboration.
- Enhanced understanding of potential information operations intended to destabilise populations and alliances.
- Motifs provide a generalised method for network analysis.

## Results

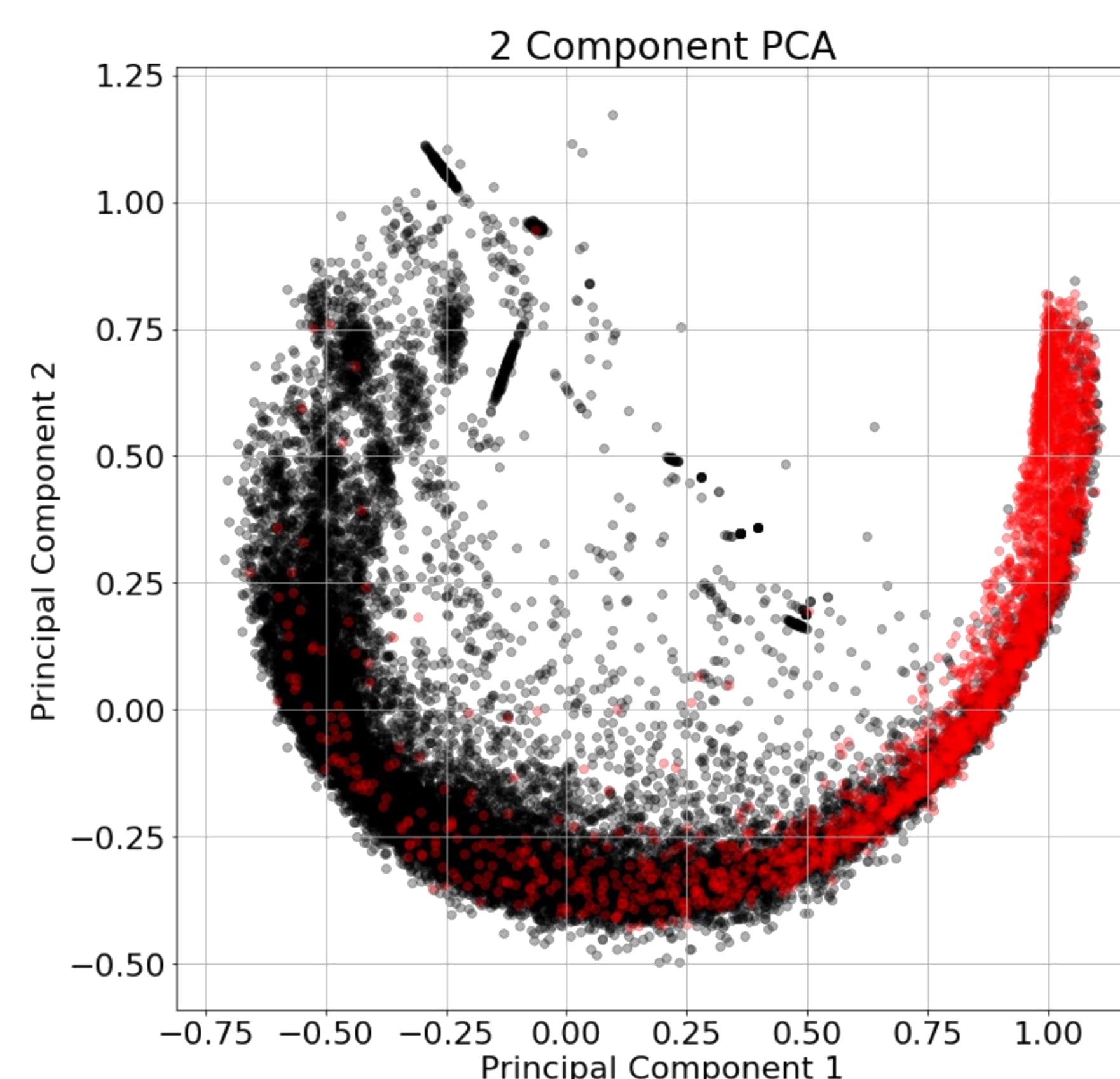
- Identifies **significant clustering of editor interactions** using PCA in two-dimensional clustering space
- Non-controversial content produces **small incremental changes**
- Controversial networks produce distinct profiles



Collection of all 16 possible isomorphic subgraphs containing three nodes.



Overlapping of all Subgraph Ratio Profiles. Controversial articles are shown in red and non-controversial in black.



PCA scatter plot repression of the 13-point feature vectors in 2D clustering space.

## Summary & Future Work

- Prediction with additional metrics
- New insights for article structure
- Small number of features to classify content
- Use of **network motifs** to better understand **behavioural dynamics**

## Publication(s) & Impact

- Ashford, James, et al. "Understanding the Signature of Controversial Wikipedia Articles through Motifs in Editor Revision Networks." *Companion Proceedings of The 2019 World Wide Web Conference*. ACM, 2019.