An Interactive Framework for Profiling News Sources
Nikhil Mehta and Dan Goldwasser

(1) Motivation
- Social media is very influential in today’s society
- But...harmful information can spread quickly
- Fake news, Politically biased content
- Important to detect it!

- News is being spread on new topics everyday, these are emerging news events
- Ex: When COVID-19 happened, a pandemic like that was never seen!
- Makes task harder!

- News often spreads in communities on social media
- Tightly connected groups of users
- Some communities are more likely to spread fake news, others not
- Identifying communities can be useful!

- But Identifying communities is hard!
- Even the best models struggle!
- Humans can help, but can’t identify everything
- We can use LLMs + Humans + Graphs to do better!

- Key Contribution: Propose an interactive approach using LLM, Graph, and Human insight to profile news sources better!

(2) Task Defn.
- News Media Profiling: Detecting the factuality and political bias of news sources
- Fake News Detection: High, Low, or Mixed Factuality
- Political Bias : Left, Right, or Center

(2) Approach: Graph
- Model data in an information graph
- Nodes: News sources, users, and articles
- Edges: Relationships between nodes
- Train graph for the source detection tasks

(3) User Communities
- Users form communities on social media
- Similar users likely have similar interests and share similar content, which has similar factuality/bias
- Idea: If we can model user preferences and thus user similarity, we can better model the content they propagate

(4) Interactive Approach
- Question: How to identify user communities?
- Humans can look at a few users and determine if they are similar
- Graph architecture models user preferences, captures community structures, but likely incorrect
- LLMs can look at text and determine textual similarity, but can’t model user preferences well
- How can we best use this?

- Approach:
  1. Cluster users in graph model -> get candidate user communities
  2. For each community, determine users the graph model is unsure about
  3. Ask humans, are these users similar?
  4. Ask LLMs to identify the same user similarity the humans identified, but on other sets of users. Text similarity detection is easier for LLMs!
  5. Connect edges between users LLM identifies as similar in the graph
  6. Now, profile sources better!

(5) Results and Discussion

RQ1: Do Interactions Help?
- Evaluate on Black Lives Matter data
- Evaluation setting: No test data, users, or topics seen at test time, hard!
- Tab 1: Human interactions lead to significant improvements
- ~33% improvement on FN, ~40% on Bias

RQ2: How are LLMs used?
- Fig 2: LLMs get few-shot examples based on human interaction
- LLMs tasked to either “accept” or “reject” new users into the human formed community

Future Work
- Building larger and better communities
- Application on other social media tasks
- General framework to use human insight to help LLMs

Table 1: Results

<table>
<thead>
<tr>
<th>Model</th>
<th>FN Acc</th>
<th>FN F1</th>
<th>Bias Acc</th>
<th>Bias F1</th>
<th># Users</th>
<th># Sources</th>
<th># Edges</th>
<th># Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline: (Mehta et al., 2022)</td>
<td>41.89</td>
<td>48.79</td>
<td>27.43</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Graph Only: High Parity 2 Communities (Comms.)</td>
<td>43.01</td>
<td>46.15</td>
<td>28.59</td>
<td>25.25</td>
<td>1,200</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Graph Only: High Parity 4 Communities (Comms.)</td>
<td>41.89</td>
<td>48.71</td>
<td>21.83</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LLM Only: 2 Commits, 2 Expansion Rounds</td>
<td>42.70</td>
<td>45.01</td>
<td>27.84</td>
<td>38.63</td>
<td>494</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LLM Only: 4 Commits, 2 Expansion Rounds</td>
<td>42.70</td>
<td>39.50</td>
<td>33.22</td>
<td>69.56</td>
<td>1,791</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LLM Only: 6 Commits, 2 Expansion Rounds</td>
<td>40.54</td>
<td>37.03</td>
<td>29.22</td>
<td>73.63</td>
<td>1,612</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LLM + Humans: 2 Commits, 2 Expansion Rounds</td>
<td>52.56</td>
<td>44.23</td>
<td>33.40</td>
<td>25.26</td>
<td>367</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>LLM + Humans: 2 Commits, 4 Expansion Rounds</td>
<td>46.36</td>
<td>46.35</td>
<td>45.13</td>
<td>72.56</td>
<td>1,087</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>LLM + Humans: 4 Commits, 2 Expansion Rounds</td>
<td>43.01</td>
<td>47.43</td>
<td>32.00</td>
<td>55.43</td>
<td>808</td>
<td>2</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>LLM + Humans: 6 Commits, 2 Expansion Rounds</td>
<td>41.34</td>
<td>48.07</td>
<td>33.91</td>
<td>82.61</td>
<td>1,696</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>