Review of the Project Presentation for group 2:

**Topic of presentation:** Online News Aggregator with Spatial Diversity.

**Group Members:** Ahmed Eldawy and Amar Ahmed

**Reviewed by:** Anuj Karpatne and Vijay Borra --Group 10.

The review is done with respect to the following elements.

**Problem Motivation:**
The motivation of the problem is not mentioned in the presentation. Though, it is not hard to understand the problem and its usefulness in real life scenario. The problem comes in the domain of applied research. The speakers talk about aggregating online news using spatial diversity. But since there are many other websites doing online news aggregator, a slide mentioning why they have to include spatial diversity, what are the features that the existing ones are missing could have been included. Considering the amount of people nowadays reading news online the idea could be very useful in terms of the societal importance.

**Problem Statement:**
The input output and the constraints are mentioned clearly in the presentation. A discussion on why the problem is hard is also mentioned.

**Challenges:**
Few of the challenges presented are
1. Limitations in the present state of art algorithms
2. The problem is NP-hard
3. Optimal substructure property might not hold

All the challenges mentioned are computational in nature.
Proposed Approach:
The talk gives the current state of art algorithms like Minack and GMC. The given algorithms do not consider the spatial diversity. The proposed approach models the above said algorithms to work with the spatial data. The proposed approach takes a greedy method of finding the locations from where the news is supposed to be picked. It may not solve the NP-hard nature of the problem but might solve the optimal substructure property.

Novel/Better:
The presentation discusses its novelty in
1. Including the spatial diversity aspect of the data
2. Improving the scalability of the GMC algorithm
3. Improving the quality of the Minack Algorithm

The novelty of spatial diversity is visible from the implementation. The other aspects of novelty and their validation is given in the validation section.

Validation:
The aspect of spatial diversity is included in the algorithm. A graph showing the performance of the modified algorithm with respect to the previous algorithms are also given. The scalability and quality are discussed with respect to the number of news items processed and the number of news items given to the user. There was a live DEMO of the application. Some of the results in the application are contrary to the results expected. Some of the items that are very far from each other are not associated to each other and some that are very close are associated with each other.
Presentation Critique:

- *Was the talk accessible to an "intelligent lay person"?*
  The talk is accessible to intelligent lay person as the presentation was full of live examples, screenshots and live demo.

- *Did the talk emphasize a central message that conveys the overall value of the work being executed?*
  The talk does emphasize on the central message. The features mentioned are driven to solve the central message.

- *Did the talk attempt to relate to the audience and showed effort in conveying key ideas clearly?*
  Yes. The talk is clear and concise and people are able to relate to that easily.

- *Was the speaker's response to questions satisfactory?*
  There were lots of questions asked. The questions were related to the implementation of the algorithm and some are comparison of the algorithm with respect to clustering algorithms. Most of them are answered.

How did the talk do on covering the 6 elements? Kindly rate each element separately and include a brief justification for each.

**The overall rating of the presentation would be 8.5/10**

The talk conveyed the six elements in the following way. The rating for the same would be:

*Problem statement:* 10/10 Given clearly. Input output and challenges are given.

*Problem motivation:* 6/10. The motivation is not mentioned. Why a need for spatial diversity. Limitations of existing aggregators could be addressed.

*Proposed approach:* 8/10 This was clearly stated using screenshots and the execution trace of the algorithms.
**Challenges:** 9/10 The challenges given are practical and real. The proposed approaches overcome few of the challenges.

**Novelty:** 9/10 The novelty is addressed clearly and the implementation seems to support the novelty. But the live demo showed some aberrations to the expected results.

**Validation methodology:** 7/10 Validation to the contribution claims are shown by comparing the two methods’ quality and scalability. As mentioned, the live demo showed some results that are not expected.

**Suggestion:**

1) Using distance measure alone for these topics could be a shortcoming. More emphasis on the content and the importance of the news in selecting it from a neighborhood can also be given.

2) Breaking news into clusters (in terms of topic they address) and applying the algorithm to each cluster could be looked at.

3) Often on major news almost all over the world the same news could be the top news but with a different wording. Care can be taken that the algorithm does not give the same news from different parts, but gives different news from different parts.

4) Often two interesting news can be correlated like a WTC bombing and Laden message from Afghanistan, or Italy winning football over Germany in the World Cup held at South Africa and Germany’s president response to it. Considering the above cases can be done. When displaying only top 5 headlines care can be taken to omit correlated items. But when displaying say 50 headlines this can be incorporated.