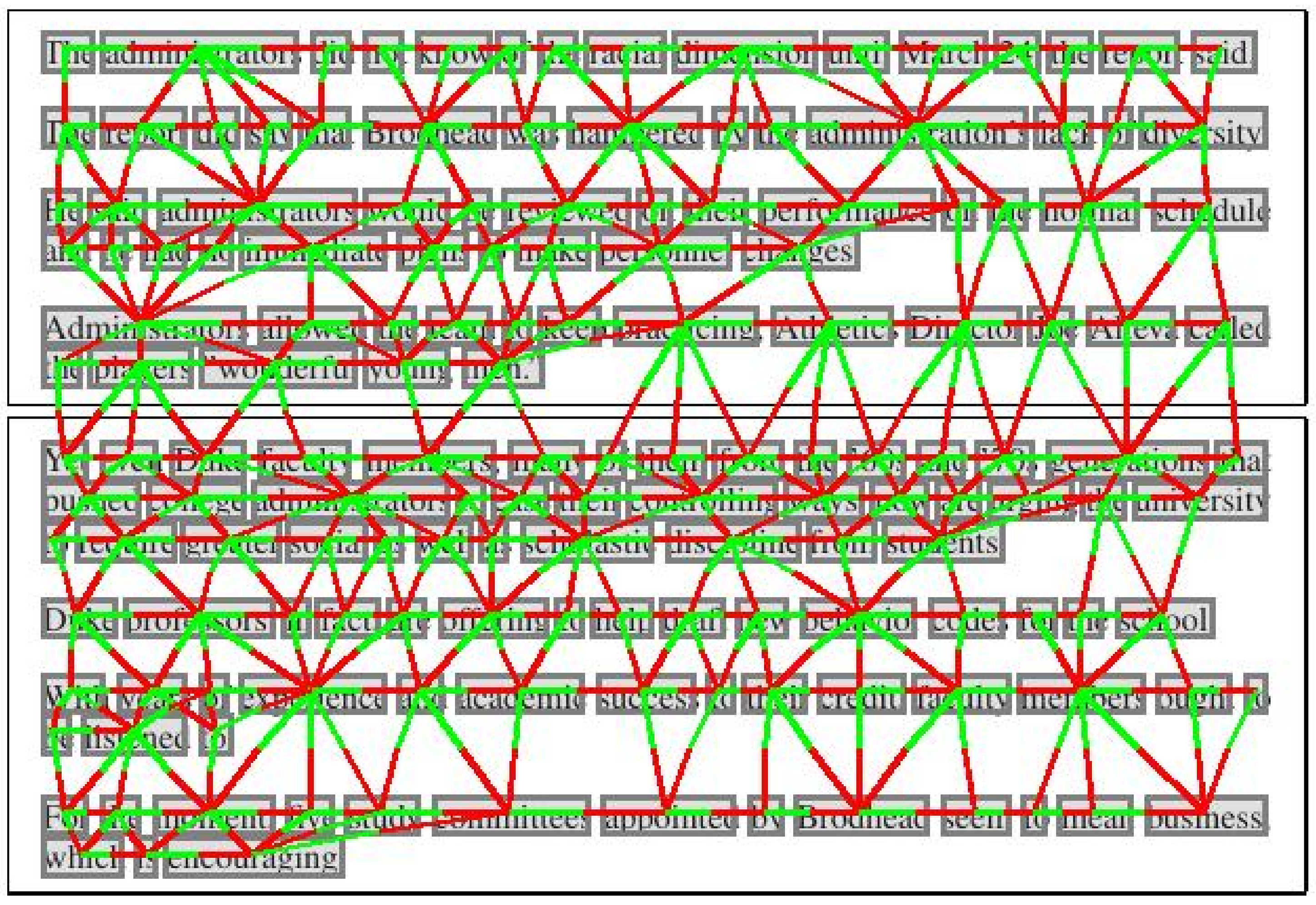

Table 3: Two topics representing a different perspective on the same event

sentences, but a cluster of sentences representing an event. Accordingly, we combine the feature vector representations of the single sentences in an event cluster into one feature vector, forming an aggregate of all their features. Although at this stage we have everything we need to infer activity times, our approach allows incorporating information from multiple articles.

3.4 Inter-Article Event Representation

To represent events over multiple articles, we suggest two methods for Inter-Article (IRA) topic modeling. The first, IRA.1, is to combine the articles and treat them as one large article. This allows processing as described in IAA, with the exception that event clusters may contain sentences from multiple articles. The second, IRA.2, builds on IAA models of single articles and uses them to construct an IRA model. The IRA.2 model is constructed over a corpus of documents containing event clusters, allowing a grouping of event clusters from multiple articles. Event clusters may now be composed of sentences describing the same event from multiple articles, thus increasing our pool of explicit temporal expressions available for inference.

3.5 Activity Time Assignment

To accurately infer activity times of all sentences, it is crucial to properly utilize the available temporal expressions in the event clusters formed in the IRA or IAA models. Our proposed inference algorithm is a starting point for further work. We use the most frequent activity time present in an event cluster as the value to assign all the sentences in that event cluster. In phase one of the algorithm we process each event cluster separately. If the majority of sentences with temporal expressions have the same activity time, then this activity time is distributed to the other sentences. If there is a tie between the number of occurrences of two activity times, both these times are distributed as the activity time to the other sentences. If there is no majority time and no tie in the event cluster, then each of the sentences with a temporal expression retains its activity time, but no information is distributed to the other sentences. Phase two of the inference algorithm reassembles the sentences back into their original articles, with most sentences now having activity times tags assigned from phase one. Sentences that remain unmarked, indicating that they were in event clusters with no majority and no tie, are assigned the majority activity time appearing in their reassembled article.

4 Empirical Evaluation

In evaluating our approach, we wanted to compare different methods of modeling events prior to performing inference.

- Method (1) IAA then IRA.2 - Creating IAA models with 20 topics for each news article, and IRA.2 models for each of the three sets of IAA models with 20, 50, and 100 topic.
- Method (2) IAA only - Creating an IAA model with 20 topics for each article
- Method (3) IRA.1 only - Creating IRA.1 model with 20 and 50 topics for each of the three sets of articles.

4.1 Results

Table 4 presents results for the three sets of articles on the six different experiments performed. Since our approach assigns activity times to all sentences, overall accuracy is measured as the total number of correct activity time assignments made out of the total number of sentences. The baseline accuracy is computed by assigning each sentence the article publication date, and because news generally describes current events, this achieves remarkably high performance.