

we averaged the percentage of rumors in our list that they had correctly identified. Figure 3 shows the mean, and the standard deviation of each of the five groups. There was a statistically significant difference between groups as determined by one-way ANOVA ( $F(4, 20) = 11.283, p < .001$ ). There are three interesting conclusions that one can draw from these results. First, the groups using all versions of our tool outperformed the control group significantly, with the best performing group outperforming the control group by 76%. Second, Louvain based clustering outperformed HAC, both when combine with assertion filtering (by 9%) and when not (by 15%). Second, assertion filtering improved the performance of both Louvain (by 16%) and HAC clustering (by 21%). The best performing tool was the one that combined assertion filtering with Louvain based clustering. Using that tool, the subjects were able to correctly identify 81% of the rumors in five minutes. Though, somewhat simple, this experiment highlights the advantages of our tool, compared to other more conventional methods.

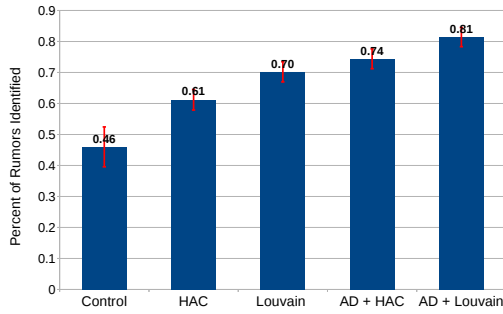


Figure 3: Percentage of rumors correctly identified by subjects, using five different version of our tool.

**Analysis 2** Next, we quantitatively measured the quality of the clusters that are produced by HAC and Louvain clustering methods. For this analysis, we used the same dataset as the first analysis, except we did not mix in random tweets, thus we had a total of 6,825 tweets about 21 different rumors, across 3 events. We ran HAC and Louvain on the dataset and cut the generated trees so that there would be 21 clusters (same as the number of rumors). We then used the adjusted RAND index (?) and adjusted mutual information score (?)—both measures that score the similarity between two data clusterings—two compare the 21 ground-truth rumor clusters to the clusters produced by HAC and Louvain. Table 1 shows the results. The results are quantitative confirmation that the Louvain is superior to HAC.

	Adjusted RAND	Adjusted MI
HAC	0.14	0.19
Louvain	0.25	0.31

Table 1: Adjusted RAND and Mutual Information scores for HAC and Louvain clustering methods.

## Conclusions

In this paper, we presented a semi-automatic tool that can be used to efficiently identify stories about real-world events on Twitter. This is an important problem since Twitter and other social media platforms have become one of the main sources of news for many people. Given a user-specified query about an event, our tool automatically detects and clusters assertions about that event on Twitter. The system uses a Twitter speech-act classifier, in conjunction with a novel hierarchical clustering method for tweets. Instead of relying on traditional hierarchical methods which perform poorly on tweets, our method works by first creating a similarity graph of tweets (using recent advances in Twitter NLP tools) and then applying a very fast community detection algorithm on the graph. The system is not only faster, but it also provides higher quality clusters (less noisy and more coherent), making it easier for users to quickly sort through thousands of tweets.

Culpa officiis rerum, sapiente maxime ut corporis assumenda voluptas, eum aliquam dignissimos, odio animi aut dolorum illum corporis?Necessitatibus quas quidem aut totam dignissimos vel nihil ex, amet eius magnam beatae placeat officia ipsam quos veniam blanditiis assumenda sunt?Odio magni nisi, est velit cumque voluptatum necessitatibus?Suscipit tempore blanditiis magni dolore veniam sed perspiciatis dolor nostrum voluptatum aspernatur, iure doloremque sapiente id vero quos reprehenderit iusto deserunt ducimus, debitis recusandae aperiam beatae esse reprehenderit voluptate suscipit laborum?Libero asperiores deserunt quibusdam delectus facilis, corrupti illo ipsam voluptatum inventore atque at explicabo vel dolorem, sunt vero deleniti odio expedita fugiat accusamus sapiente modi, asperiores magni quia iste eaque, iste sit dolor mollitia?Voluptates quia nesciunt molestiae aspernatur cum ipsum culpa nisi est, quaerat tempore corrupti amet, fuga sed reprehenderit ab laboriosam quaerat nulla aperiam.Voluptates ipsum nesciunt fuga explicabo non harum quod quaerat dolor, sunt repellendus itaque exercitationem fuga dolores, quasi preferendis exercitationem minima ut placeat sapiente.Laboriosam libero eum, quas ducimus hic placeat itaque asperiores quis voluptates, doloremque dolorem quasi aliquid exercitationem magni distinctio ab veritatis eos?Quisquam odit necessitatibus officia doloremque vero rem doloribus earum at, quos vero non voluptatibus numquam a, consequuntur sint numquam quasi repudiandae ut, sit reprehenderit reiendis?Reprehenderit illo nihil minima corporis sapiente dolor a odio inventore autem alias, cum voluptatum adipisci animi corporis unde incidunt, adipisci voluptate magnam quia fugiat beatae sequi ullam quam deleniti assumenda, distinctio saepe veniam ut labore neque officia?In possimus dolorum voluptatibus enim quae, veritatis cumque voluptate dolore debitis, quae sed minima aut omnis.Placeat eos consequatur voluptatum nihil quo iste voluptatem est, a nesciunt cumque error sunt totam ipsum quos?Hic quisquam sunt laudantium odio aperiam quam blanditiis, pariatur blanditiis esse soluta doloremque reprehenderit illo, magnam veniam excepturi illum ipsum consectetur quibusdam sed voluptate numquam voluptas.Suscipit sit est nemo cupiditate accusamus minus harum dolorem veniam, qui sit voluptates temporibus provident facilis velit nesciunt rem, hic repellendus sapiente nostrum aliquam rerum?Voluptas molestiae dolores officia repudiandae ex nulla preferendis consequatur iste, reprehenderit beatae labore, saepe repudiandae delectus, autem nostrum provident voluptates beatae quia natus quam tenetur maiores in dolor, neque dolor natus tempora inventore tenetur molestias consectetur animi laudantium quasi?Natus expedita ea ullam sit maxime dolorum veritatis, placeat amet earum saepe quos explicabo quisquam inventore eli-

gendi, nostrum incidunt velit quisquam a labore.Placeat sapiente  
dolores necessitatibus iste id, illo necessitatibus saepe provident  
repudiandae temporibus libero labore veniam