Supplementary Information

1 The Geographical Boundaries of Different Cities

For each event, we consider users who have frequently posted tweets around the event location prior to the event occurrence. We collected the tweets from Twitter API by specifying the geographical boundaries to identify the local users in each city. The boundaries are manually determined based on the cities' information available on the websites of the city governments, Google Maps, and Wikipedia. Here are the boundaries for each city we used:

city	city boundaries(longitude, latitude)
Paris	(2.083, 48.722), (2.613, 48.984),(2.083, 48.984), (2.613, 48.722)
New York	(-74.257, 40.559), (-73.762, 40.849), (-74.257, 40.849), (-73.762, 40.559)
London	(-0.621, 51.234), (0.299, 51.716), (-0.621, 51.716), (0.299, 51.234)
Brussels	(4.216, 50.729), (4.528, 50.944), (4.216, 50.944), (4.528, 50.729)
Orlando	(-81.656, 28.334), (-81.206, 28.701), (-81.656, 28.701), (-81.206, 28.334)
San Bernardino	(-117.359, 34.086), (-117.202, 34.166), (-117.359, 34.166), (-117.202, 34.086)

2 The Choice of Time Window

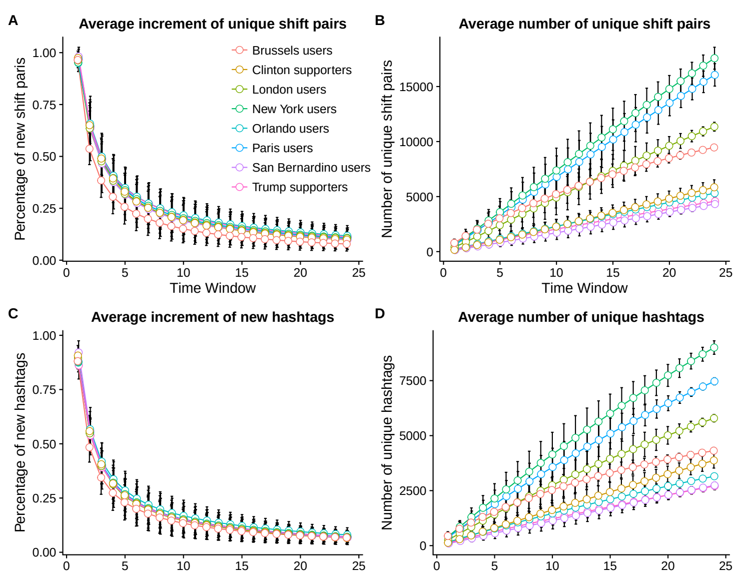


Figure S1: **The percentage of new shift pairs/nodes and the number of unique shift paris/nodes for different time window.**

A proper time window is an important parameter in building the attention graphs. The time window is determined by both the tweet volume and the hashtags adoption frequency of different user sets. If the time window is too short, the attention graphs tend to be too sparse; on the other hand, if the time window is too long, the differences between consecutive attention graphs become neglectable – both lead to non-informative graphs for extracting interesting or meaningful patterns.

As we defined in the paper, for an attention shift event of a user u is a transition between two of u's consecutive hashtag adoption events, denoted as $s = (\langle u, h_j, p_j\rangle | \langle u, h_i, p_i\rangle)$, where $h_i \neq h_j$, $p_i < p_j$, and there is no adoption event (u, h_m, p_m) s.t. $h_i \neq h_m$, $p_i < p_m < p_j$. The hashtag pair (h_i, h_j) is the shift pair induced this shift event. Let $S_{t,n}$ denotes all the unique shift pairs in $[t-n, t]$ and $H_{t,n}$ denotes all the unique hashtags used by any user in $[t-n, t]$, where n is the time window.

We used the week prior to the event week to study the proportion of new information induced by different time window. When the time window $n \in [1, 24]$ is fixed, for every hour t in the prior event week, we computed the increment of new shift pairs between $S_{t,n}$ and $S_{t-1,n}$, defined as $|\{(h_i, h_j)|(h_i, h_j) \in S_{t,n}, (h_i, h_j) \notin S_{t-1,n}\}|/|S_{t,n}|$. Then we took the average of all the results to get the average increment of unique shift pairs as shown in Fig. S1 A.

Similarly, with the increment of new hashtags between $H_{t,n}$ and $H_{t-1,n}$ defined as $|\{h_i|h_i \in H_{t,n}, h_i \notin H_{t-1,n}\}|/|H_{t,n}|$, we can get the average increment of unique hashtags as shown in Fig. S1 C.

As we can see from Fig. S1 A and Fig. S1 C, the proportion of new information between two consecutive hours rapidly decreases as time window increases. We use the elbow area where $n \in [4, 8]$ to determine the optimal time window. Further increasing the time window will make the consecutive attention graphs comparable and computationally cost-expensive. The unique shift pairs are possible edges and the unique hashtags are possible nodes in the attention graph. In Fig. S1 B and Fig. S1 D, we reported the average number of unique shift pairs and the average number of hahstags for different time windows. The increment speed for the number of possible edges and possible nodes are different across datasets due to the differences in user set size and user activeness. For the sake of simplicity, we used two time windows in the paper, we chose a shorter time window($n = 4$) for the ones whose potential edges and nodes increase rapidly with larger time window (the Paris, Brussels, New York and London datasets). We used $n = 8$ for the remaining four datasets.

3 Tweet with Hashtags Over Time

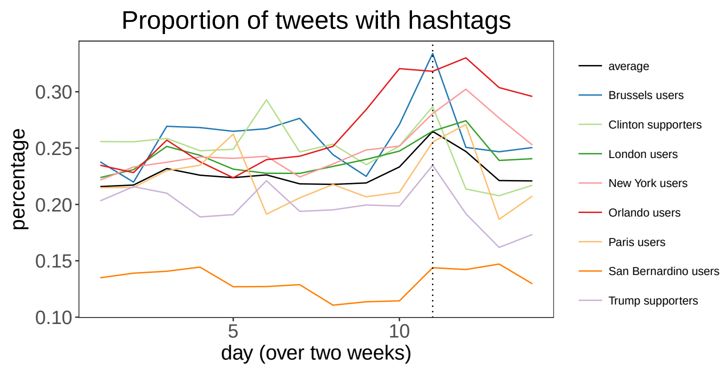
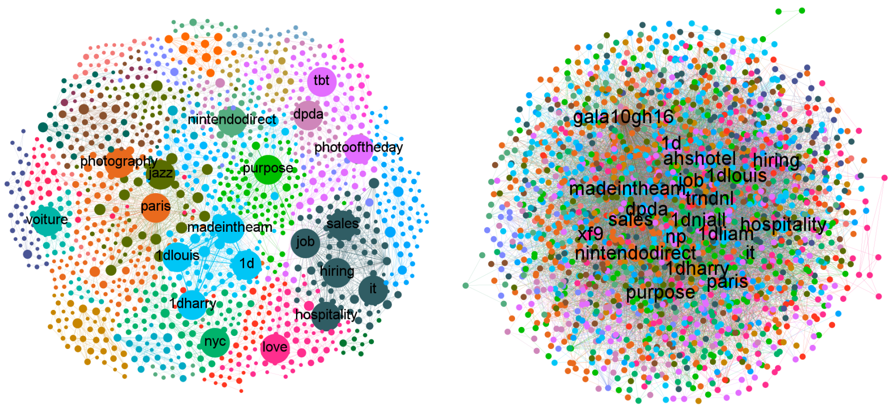


Figure S2: **The proportion of tweets containing hashtags in two weeks.** The dotted line indicates the event day, the black line is the average of the eight datasets.

Here we show the percentage of tweets that contain hashtags in two weeks: the event week with the event day centered in the middle and the week prior to the event week. From Fig. S2, we can find users tend to adopt more hashtags when a large-scale event strikes. The ratio of tweets containing hashtags is 22% on average before the event day, the the use of hashtags is more intense during an event (on average 27% in our dataset).

4 Visualization of the Original Attention Graph and the Null Model

Here we visualize an attention graph in the normal time and its corresponding randomised version generated from the null model as described in the paper. As we can see from Fig. S3, although some of the populr hashtags are similar in the two graphs, the randomised attention graph is much less structured. In the original attention graph, the popular hashtags form topical clusters and are more localized, which is not seen in the randomized version. The original attention graph has a modularity score of 0.82 whereas the randomised attention graph only has 0.46.



(a) The original attention graph.

(b) The randomised version.

Figure S3: The attention graph for the Paris users from 00:00:00 11/12/2015 to 04:00:00 11/12/2015 GMT and the corresponding randomised version generated from the null model. In each graph, we label the top 20 hashtags based on weighted degree.

5 In Comparison with the Null Model

Here we show the results of the collective attention measured by the original attention graph and the null model for each single data set. For each event, the event day is centered in the middle.

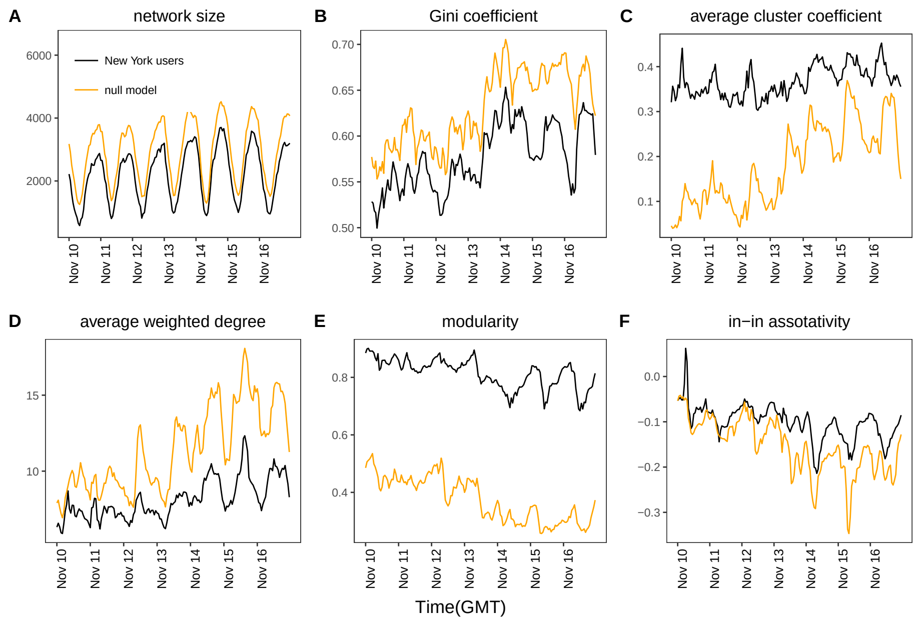


Figure S4: New York users.

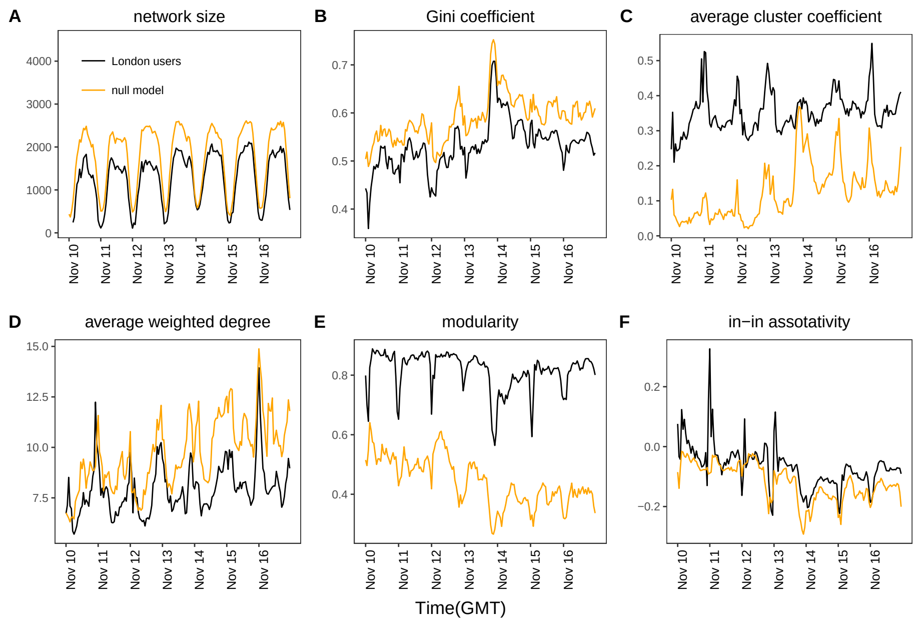


Figure S5: London users.

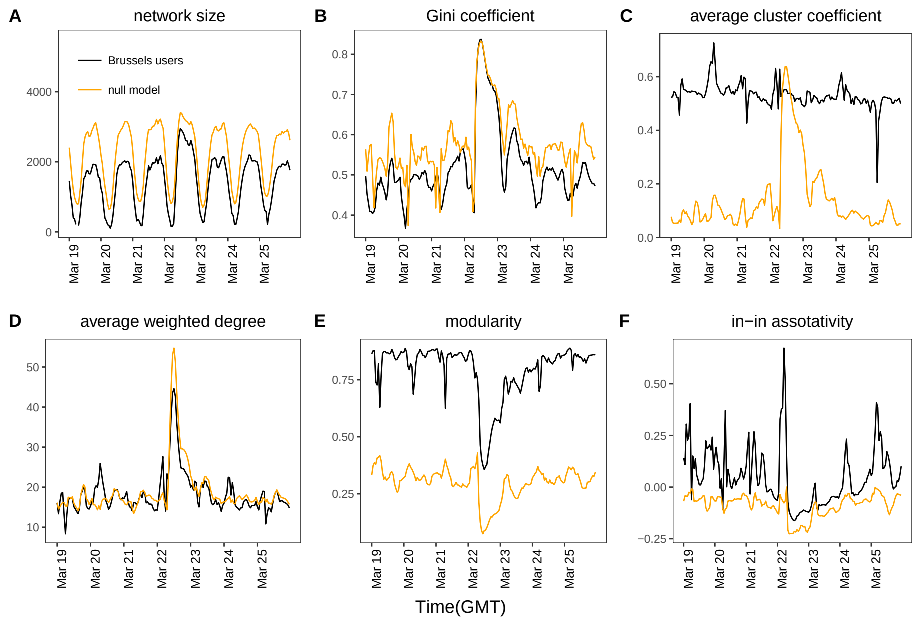


Figure S6: Brussels users.

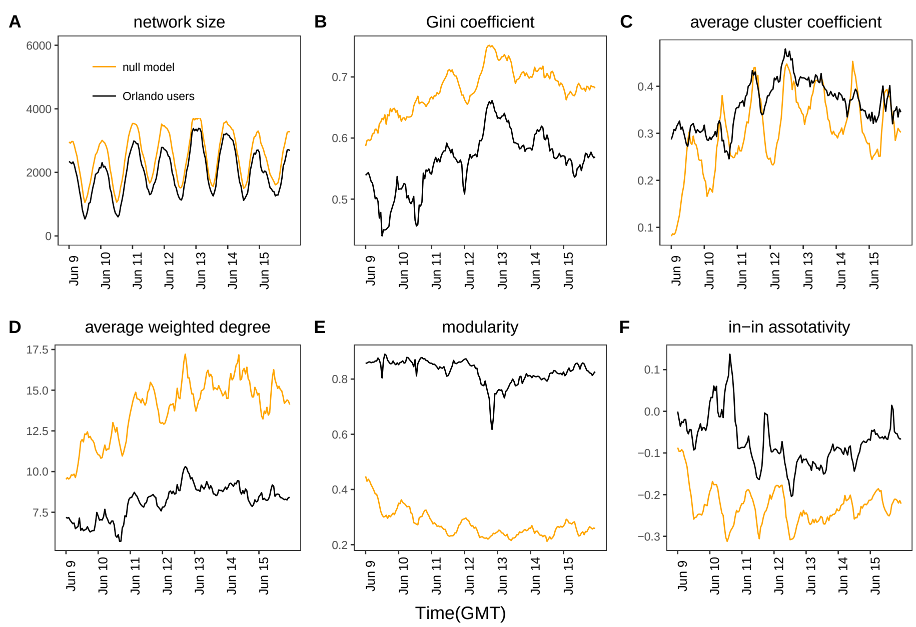


Figure S7: Orlando users.

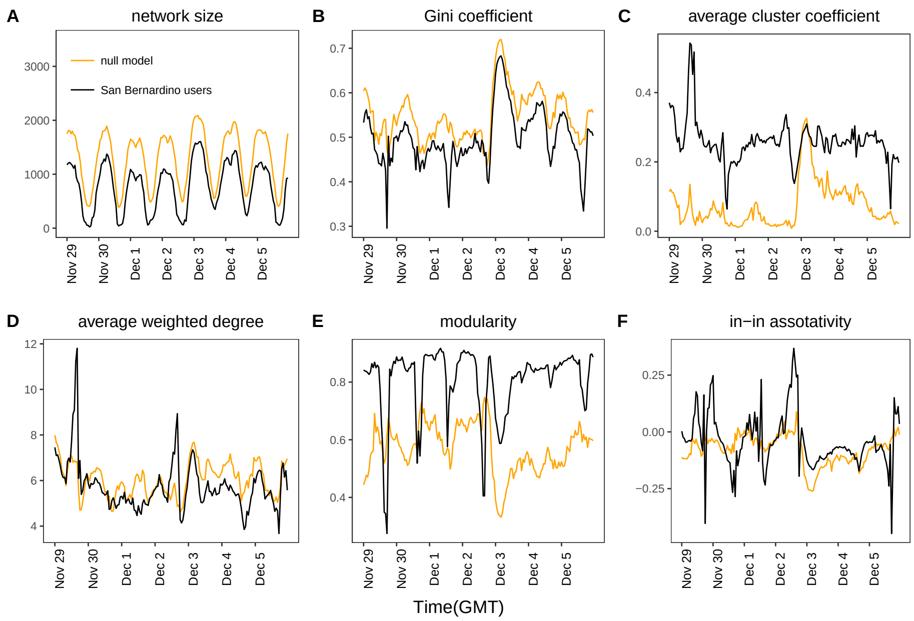


Figure S8: San Bernardino users.

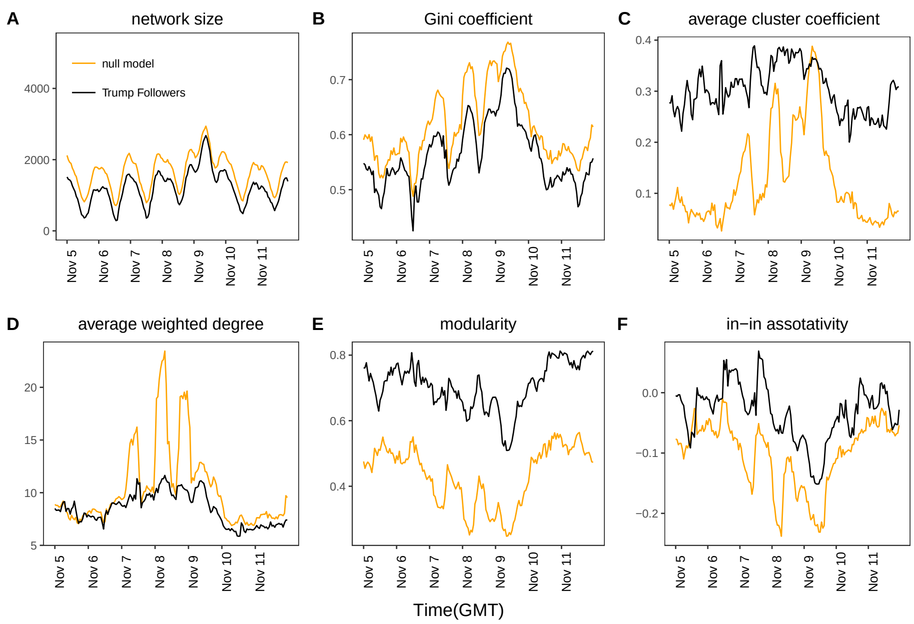


Figure S9: Trump followers.

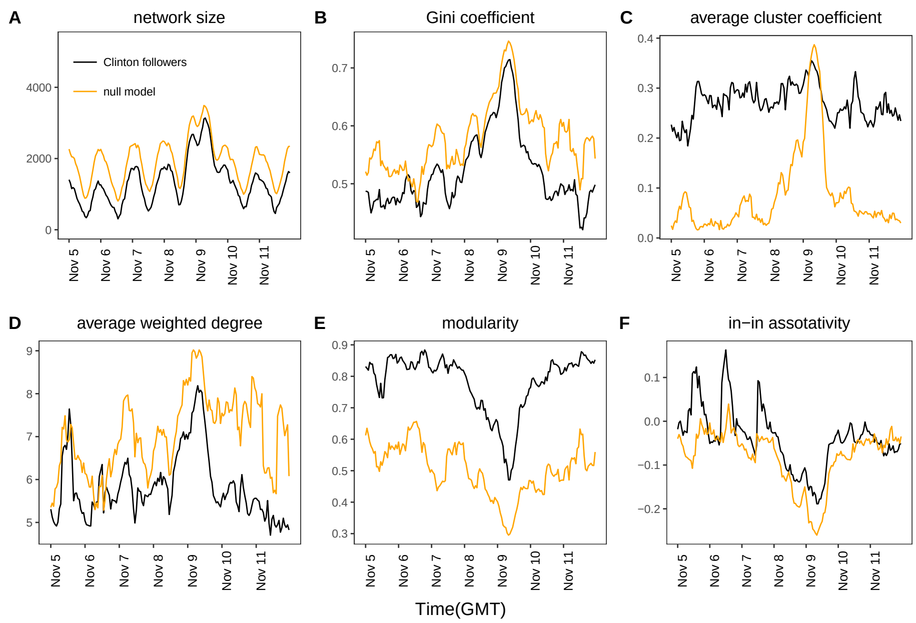


Figure S10: Clinton followers.