Multidimensional and Multilevel Analysis to Find and Describe Irregular Patterns in Political Communication on Twitter

Audrey Wilmet[1] and Robin Lamarche-Perrin[1,2]

Audrey.Wilmet@lip6.fr, Robin.Lamarche-Perrin@lip6.fr
[1]Sorbonne Universités, UPMC, Université Paris 6, CNRS, LIP6
[2]Institut des Systèmes Complexes Paris Île de France, CNRS

Session : Multilevel Network Perspectives in and around Organizations : Theory, Structure and Dynamics.

Keywords : Politics, Twitter, anomaly detection, multidimensional data aggregation, spatio-temporal network analysis.

Twitter is now an integral part of means of communication used by political leaders to disseminate information to the public. A politician may use it sporadically to merely broadcast to his followers or on the contrary employ it regularly and tweet at strategic moments. Likewise, their followers may be occasional spreaders or real online activists retweeting primarily a particular political figure. The complex processes formed by interactions between users through a retweet may result in a media event whose influence goes beyond Twitter framework. We propose a multidimensional and multilevel analysis method to describe structural and temporal relational patterns in this retweet network as well as to find unexpected behaviors related to this *political strategy* of communication.

The retweet temporal network we use has been obtained from observing the Twitter accounts of nearly $3,500$ political actors such as individuals, organizations and institutions. In this network, two users are connected at time t if one of the two – the spreader – has retweeted a tweet of the other – the author – at time t. Our method consists firstly in evaluating the *quantity of interactions* between an author and a spreader during a given time period. After this step, we have access to *local information*: the finest scale at which we can observe interactions. One then uses *data aggregation* to obtain the total quantity of interactions of an entity such as an author, a spreader, an hour or a couple obtained by combining those three dimensions (marginal values). This step gives us access to *global information* that can be used to provide more context to local data.

Afterwards, in order to *detect irregularities* in users and temporal behaviors we compare the previously obtained quantities of interactions between them. Here again, we propose to decompose comparison in multiple levels. For instance, one can look at the quantities of interactions by hours, comparing them all at global scale and find an hour with unexpected activity level compared to all the others. Then, one can look at the couples (author, hour), comparing them at local scale after having fixed the hour previously found abnormal, i.e.

comparing authors'activities on this hour, and finally find an author with unexpected activity level. Hence, while *global comparison* consists in comparing the quantity of interactions of an entity to the quantities of interactions obtained by *all* other entities of the same type, *local comparison* consists of fixing a variable in an entity and comparing its value to all others obtained by varying the remaining variable(s).

Our method enable us to explore all the different levels of interactions and to find outliers according to the normality that we define by choosing the context to which we compare the quantities of interactions. We show that, by varying information and comparison granularities, one can build interesting insights regarding the political organization on Twitter. Indeed, first results bring out political leaders and online activists among Twitter users, political meetings, relationship between a spreader and a group of authors, and so on.

This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 732942 (ODYCCEUS).

References

[1] C. Grasland, R. Lamarche-Perrin, B. Loveluck, and H. Pecout. L'agenda géomédiatique international : analyse multidimensionnelle des flux d'actualité. *L'Espace Géographique*, 2016.

[2] M. Gupta, J. Gao, C. C. Aggarwal and J. Han. Outlier detection for temporal data : A survey. *IEEE Transactions on Knowledge and Data Engineering, 26(9), 2250-2267*, 2014.

[3] V. Chandola, B. Arindam, and K. Vipin. Anomaly detection : A survey. *ACM computing surveys (CSUR) 41.3 : 15*, 2009.