

Discovering Bilateral and Multilateral Causal Events in GDELT

Lei Jiang ^{*} and Fan Mai

¹ Tapjoy Inc., San Francisco, CA 94104

² Department of Sociology, University of Virginia
Charlottesville, VA 22093 USA

Abstract. GDELT is one of the richest sources of event data and the hidden information could be potentially of tremendous value for its real-world reflections. While the majority of past research on these data sets puts emphasis on featured event series alone, such as Arab spring and violence in Afghanistan, this paper is focused on exploiting the causality between multiple types of events in different countries. In other words, for example, the actual impact of an event would result in responses from other parties, or even an escalation that more events will come thereafter. With the causal link in real-world context, time series analysis is performed on GDELT data to discover such causal links: we start from an exploration of bilateral and multilateral events in certain countries, where a strong correlation is shown; then Granger test is systematically conducted in the large-scale data set to help to identify causality in a rigorous manner. Further, with the addition of data in related parties, a scenario of prediction task is tested to show enhanced predictability as a validation of this methodology.

Keywords: GDELT, Granger causality, bilateral events, time series analysis, prediction

^{*} E-mail:lionelchange@gmail.com