Results in Physics 6 (2016) 414–419

Contents lists available at ScienceDirect

Results in Physics

journal homepage: www.journals.elsevier.com/results-in-physics

Association between volume and momentum of online searches and real-world collective unrest

Hong Qi [a], Pedro Manrique [a], Daniela Johnson [b], Elvira Restrepo [c], Neil F. Johnson [a,*]

[a] Department of Physics, University of Miami, Coral Gables, FL 33126, USA
[b] Department of Government, Harvard University, Cambridge, MA 02138, USA
[c] Department of Geography, University of Miami, Coral Gables, FL 33124, USA

ARTICLE INFO

Article history:
Received 25 May 2016
Accepted 22 July 2016
Available online 29 July 2016

Keywords:
Collective phenomena
Complexity
Social media
Transitions

ABSTRACT

A fundamental idea from physics is that macroscopic transitions can occur as a result of an escalation in the correlated activity of a many-body system's constituent particles. Here we apply this idea in an interdisciplinary setting, whereby the particles are individuals, their correlated activity involves online search activity surrounding the topics of social unrest, and the macroscopic phenomenon being measured are real-world protests. Our empirical study covers countries in Latin America during 2011–2014 using datasets assembled from multiple sources by subject matter experts. We find specifically that the volume and momentum of searches on Google Trends surrounding mass protest language, can detect – and may even pre-empt – the macroscopic on-street activity. Not only can this simple open-source solution prove an invaluable aid for monitoring civil order, our study serves to strengthen the increasing literature in the physics community aimed at understanding the collective dynamics of interacting populations of living objects across the life sciences.

Introduction

Physics has a long tradition of interest in understanding the collective dynamics of many-body, interacting systems [1]. More recently, this has developed into an interest in systems that involve active particles – and in particular, living objects such as humans [2]. In particular, there has been much interest in how macroscopic phenomena such as civil unrest, for example, emerge from a system of interacting humans through interactions in the real world or online [2]. The 2011 London riots are a good example: Davies, Fry, Wilson and Bishop [3] have given a detailed analysis of spatial development of the disorder, which can be used to examine the effect of varying policing arrangements. Given the potential by-products of rioting, looting, arson and inter-personal violence, improving our understanding of when civil unrest activity is beginning to take place – or ideally, will take place – is a highly desirable goal both in practical and scientific terms.

Inspired by the idea from physics that the position and momentum of an object, or equivalently of the center of mass of a collection of objects, can indicate the future evolution of a system's dynamics, we here investigate whether similar measures applied to collections of humans can also serve as an indicator of future collective behaviors. This is of course a hard goal to achieve, given the influence of noise and external factors in human systems, and so we see this work as a step in this direction rather than outright proof – however, the results do seem promising. Specifically, we investigate whether the volume and momenta of online searches from open-source data can be used to detect, or even pre-empt, on-street civil unrest activity without the need for any information about likely individual participants, individual-level communications or individual Internet accounts. The issue of anonymity is an important one given current societal concerns about privacy and state agency snooping. It is well known that just as social media might be used in the coordination of such activity, it offers the potential of a digital footprint – particularly in terms of private communications through, for example, applications such as WhatsApp. However information from private accounts and conversations ordinarily cannot be obtained by policing authorities without prior permission, and not in real-time. More importantly, the use of information from individual accounts raises sensitive political issues concerning privacy. Hence it would be far more desirable if an alert system could be developed based on open source information – and in particular, built around the quantity of traffic but without any need to access information about individual identities or behaviors.

Our results suggest that open-source activity surrounding Google Trends can indeed act to detect the onset of bursts of civil

* Corresponding author.

http://dx.doi.org/10.1016/j.rinp.2016.07.008
2211-3797/

unrest, and may also pre-empt the onset by a day or more. Instead of just the volume of particular keyword searches being important, we find that the combination of the volume and the rate-of-change of the volume (i.e. momentum) acts as the alerting signal. This can be understood by the fact that civil unrest is a time-varying coordinated interaction between individuals, groups or populations, with built-in feedback between events in the past and present [4]. We are cognizant of the fact that in social science, questions about the conditions under which protest, civil wars, terrorism, and other forms of violence occur are among the more persistent in the field [4–21]. However we do not need to concern ourselves here with such questions about root cause or individual motivations for participating in such unrest and civil disobedience. Instead our paper focuses on the practical scientific relationship between online and on-street activity. A limitation of our paper is that we does not show results for all possible sources of such online social interaction. It may turn out that one medium, e.g. Google Trends, is more or less capable than another. Instead our goal is a proof-of-principle that a freely available instrument such as Google Trends can be a sufficient medium for detecting and also potentially pre-empting civil unrest. We stress that we are not looking at specific mentions of planned future mob meetings embedded within online conversations, e.g. a Tweet containing a particular future date on which a civil unrest event has been planned at a particular location. Instead we want to see if the dynamical swell of online, open-source and user-anonymous activity surrounding particular keywords acts as a detector and even pre-cursor. To set our study in context, the paper begins with a background of the literature on such civil unrest, before moving to the method employed here concerning online activity. Then we present the results before finishing with the conclusions.

Background

Since this is a heavily interdisciplinary topic that stretches well beyond the background of many in the physics community, we first present a review of the wider literature surrounding the study of civil unrest. Studies of civil unrest have historically evolved from the viewpoint that protests represent 'irrational behavior', to the study of individuals within these collective actions – and ultimately causal analysis. With a few exceptions (see Ref. [3] and references therein), the existing literature on civil unrest and associated crime necessarily pre-dates the invention of online social media and the global ubiquity of smartphones. In terms of quantitative models of civil unrest, Lichbach's 1992 paper [4] provides a survey of the multiple studies that have been offered. Most of the theoretical methodology surrounds game theory and stochastic models. Siegel [5] attempts to provide a model involving social networks in which there is interdependent decision making, allowing for heterogeneity by including different motivations for individuals to participate and four different types of networks, i.e. small world, village, opinion leader and hierarchy. The main finding is that a small-world network – which is precisely that favored by online social media – is more likely to yield participation or activity that then permeates through the interconnected population. Macy [6] removes the assumption that individuals act through rational choice, and instead implements a stochastic-learning approach that leads to a network chain reaction. In this way, he incorporates evolutionary tendencies in the threshold distribution and allows decisions to join an on-street protest to be affected by other group members. Rahmandad and Sterman [7] lay out the pros and cons of utilizing agent-based versus differential equation modeling, finding a preference for agent-based modeling for situations of diffusion where agents are heterogeneous since this allows for specific network analysis. Work has also been developed toward analyzing the role of the general media in civil unrest [8] following the line of contentious politics and beyond [9–12]. Rainie et al. [9] look at networks in social operations while Centola and Macy [10] focus on the potential role of loosely connected individuals in amplifying the spread of unrest. In Ref. [11], the authors look at broadcasters and hidden influentials in online protest diffusion. Onnela and Reed-Tsochas present a mechanism for the spontaneous emergence of social influence in online systems.

Morales et al. [13] follow a specific social network during a protest in Venezuela on Twitter. They observe a growth in activity during the protest – a result that we also observe, but with the additional feature in our case that the data are entirely open-source, coming from the numbers supplied by Google Trends as opposed to actively querying Tweets. Back in 1992, Singer [14] carried out a landmark study of the role of mass media and communication processes in the Detroit riot of 1967 using public opinion survey research. Importantly, Singer found that contagion of riots can occur through word-of-mouth and also mass broadcasting via television which acted to change public opinion about the riots. He claimed that different media perform different functions in transmitting information about the riot. Most relevant to our study is the fact that Singer found that individuals just finding out about the riot or its location was enough to give rise to their subsequent action. In this sense, the modern-day act of going to Google to search for words like 'riot' should arguably play the same role – and does indeed do so according to our findings.

Other related works concerning strikes [15,16] compared the spreading of collective mobilization to forest fires, focusing on the cases of Chicago and Paris in the late nineteenth century, and also the role of positive feedback during the American Strike Wave of 1886. Meanwhile, Braha [17] looked at global civil unrest as a generalized epidemic phenomenon, finding universality emerging because social-unrest contagion is governed by the same mechanisms in spite of differences in culture. This is particularly relevant to our work, since it suggests that the core features should be independent of city or country – a suggestion that is indeed confirmed by our findings. This narrative also feeds into the extreme of insurgencies and terrorism which, while of course different in substance and execution, also involve the mobilization of actors who might otherwise be passive, and potential issues such as online self-radicalization as a result of searching for, and finding, particular materials on the Internet [18–21]. Related to this is the issue of how underground organizations manage to develop [22,23]. Ref. [22] shows that the level of organization in different dark networks such as delinquent groups and gangs, is an important variable in determining the relationship between delinquency and membership and delinquency. In Ref. [23], Gambetta imports insights from the field of semiotics and microeconomics to analyze the cost of such signaling – which can now be taken to include online communications. As such, using Google Trends is indeed a far less costly venture for a potential protestor than sending a personal Tweet to ask a friend for specific information.

Methods

Our methodological approach to testing out the relationship between online search activity and the appearance of on-street protest mobs, involved three steps: (1) building a detailed dataset of actual on-street protest events and hence unexpected mob appearances within distinct countries in a single continent; (2) examining different search engines and search words, to uncover the variation in volume over time of particular keywords; (3) comparison of these results with the actual events, in order to determine the optimal choice of search engine, word combination, and measure of search volume and its derivatives.