Journal of Complex Networks Advance Access published January 27, 2014

Journal of Complex Networks (2014) Page 1 of 12
doi:10.1093/comnet/cnt022

Critical transitions in social network activity

Christian Kuehn[†,‡]
Institute for Analysis and Scientific Computing, Vienna University of Technology, 1040 Vienna, Austria
[†]Corresponding author. Email: ck274@cornell.edu [‡]*Equal contribution*

Erik A. Martens[‡]
Department of Biomedical Sciences, Copenhagen University, Blegdamsvej 3, 2200 Copenhagen, Denmark and Max Planck Institute for Dynamics and Self-Organization, 37077 Göttingen, Germany
[‡]*Equal contribution*

and

Daniel M. Romero
School of Information, University of Michigan, Ann Arbor, MI 48109, USA and Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208, USA

Edited by: Mason Porter

[Received on 13 August 2013; accepted on 28 December 2013]

A large variety of complex systems in ecology, climate science, biomedicine and engineering have been observed to exhibit tipping points, where the dynamical state of the system abruptly changes. For example, such critical transitions may result in the sudden change of ecological environments and climate conditions. Data and models suggest that detectable warning signs may precede some of these drastic events. This view is also corroborated by abstract mathematical theory for generic bifurcations in stochastic multi-scale systems. Whether such stochastic scaling laws used as warning signs for *a priori unknown* events in society are present in social networks is an exciting open problem, to which at present only highly speculative answers can be given. Here, we instead provide a first step towards tackling a simpler question by focusing on *a priori known* events and analyse a social media data set with a focus on classical variance and autocorrelation warning signs. Our results thus pertain to one absolutely fundamental question: Can the stochastic warning signs known from other areas also be detected in large-scale social media data? We answer this question affirmatively as we find that several *a priori* known events are preceded by variance and autocorrelation growth. Our findings thus clearly establish the necessary starting point to further investigate the relationship between abstract mathematical theory and various classes of critical transitions in social networks.

Keywords: tipping point; critical transition; social networks; scaling law; word frequency; warning signs.

1. Introduction

Can sudden changes in society be anticipated by monitoring social media activity? This problem has recently gained considerable attention and serves as the background motivation for our work. It has been suggested [1,2] in the media that certain societal-scale oppositional movement may be preceded, and even be triggered, by social media activity and that those events could be viewed as 'tipping points' [3]. For example, the events in Egypt during 2011 have been the focus of recent research [4–6] with the shared conjecture that 'Tahrir Square was a foreseeable surprise' [7]. To address this conjecture from a theoretical modelling perspective would require a deep understanding of how social media impacts collective action [8] and how cyber-collective movements are formed [9]. A conclusive study to confirm the conjecture that the *a priori* unknown revolution in Egypt in 2011 was foreseeable is far beyond the reach of current research. Hence, since we cannot solve this difficult problem directly, we must reduce it to several elementary (and thus more tractable) problems concerning complex network theory. One such necessary first step towards solving this larger problem is to ask the following question: can well-known warning signs preceding drastic events in social media be found at all? In this study, we answer this question by analysing social media data for warning signs in the context of well-defined *a priori* known events such as special public holidays.

In particular, our approach establishes a potential link between social media analysis and the recent theory of warning signs for critical transitions (or tipping points) [10,11]. Similar approaches have received major recent attention in ecology, where warning signs were detected in experimental [12–14] and field data [15,16]. A critical transition may informally be defined as a rapid and drastic change of a time-dependent dynamical system; for more precise definitions see [17,18]. Warning signs for critical transitions have been investigated intensively in ecological models during the last decade [19,20]. Similar results have also been obtained in the context of climate science [21–23], biomedical applications [24,25], engineering [26] and epidemiology [17,27]. These studies led to the conjecture that there are some warning signs which are generic [10] for large classes of natural systems. From a mathematical perspective, this conjecture can be made precise for transitions near certain bifurcation points; see [11,17] and Appendix A. Two of the most classical warning signs are rising variance and rising auto-correlation before a critical transition [10,28]. The theory behind these warning signs is described in more detail in Appendix A. The basic idea is that if a drastic change is induced by a critical (bifurcation) point, then the underlying deterministic dynamics becomes less stable. Hence, the noisy fluctuations become more dominant as the decay rate decreases close to the critical transition. As a result, (a) the variance in the signal increases, due to the stronger fluctuations and (b) the system's state memory (i.e., auto-correlation) increases, due to smaller deterministic contraction onto a single state [10,11]. It can be shown that both warning signs are related via a suitable fluctuation–dissipation relation [29].

For social networks, the situation is much less developed. Although the notion of 'tipping' is somewhat familiar in sociological contexts [30,31], work on detailed statistical analysis of warnings in social networks from a dynamical systems perspective is very sparse. One approach, which is related and complementary to the results presented in this paper, is the recent work by Slater [32], where a data set of approximately 11,000 blog posts is analysed and tipping is defined using a sentiment score. Based on this score, two special points were identified and warnings were computed. In this paper, we analyse a large-scale social media data set and focus on clearly defined events, which are well localized in time.

2. Results

The data set we analysed consists of messages communicated publicly via Twitter. In this context, a message is also called a tweet. The messages collected account for about 20–30% of all the data tweeted world-wide in the time period from 1 June 2009 to 31 December 2009, amounting to 476,547,774 tweets, or on average over the time period 92,767 tweets per hour [33]; more details about the data set can be found in Appendix B.

We extract time series by counting word frequencies. Twitter users frequently attach hashtags to particular events or topics. Here we focus on hashtags such as `#halloween`. For the words we consider, we remove all white spaces and transform strings to lower case, e.g., `#halloween` equals

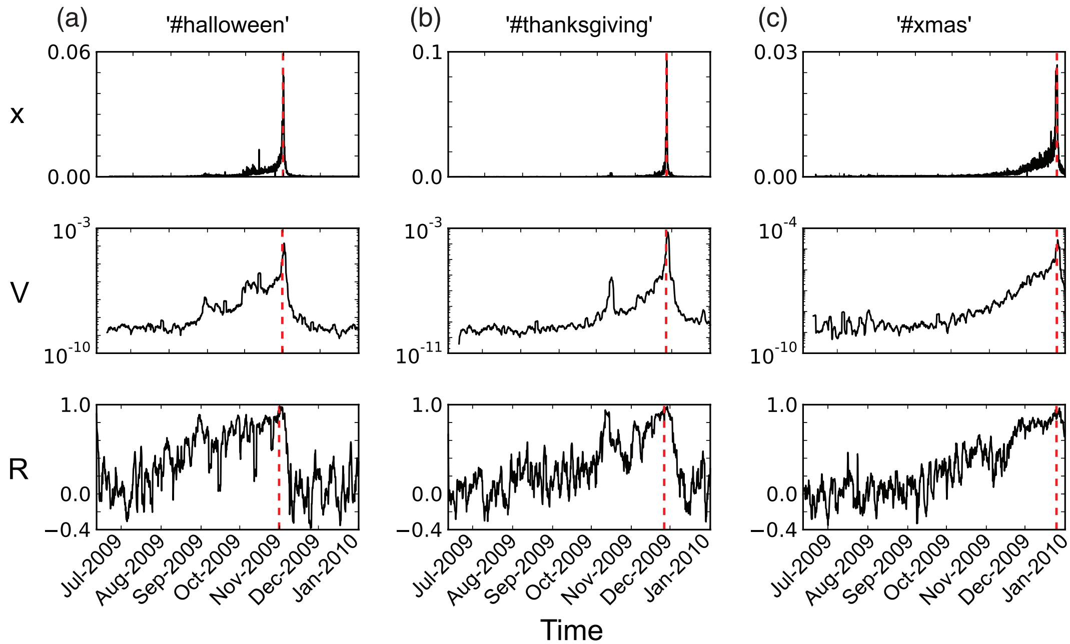

FIG. 1. Hashtag frequency time-series reveal signatures of critical transitions. (a) #halloween (max. 6019 tweets per hour) (b) #thanksgiving (max. 10157 tweets per hour) (c) #xmas (max. 3501 tweets per hour). Columns from top to bottom display normalized word frequency x, variance V and lag-1 autocorrelation R calculated from the time series of word frequency. The red vertical dashed line indicates the *a priori* known event date.

`#Halloween`. Subsequently, linear trends were removed from the resulting time series; for more details on the computation see Appendix B.

We have already pointed out that it is difficult to analyse *a priori unknown events*, such as the revolution in Egypt in 2011, using social media data. For example, it is unclear which specific hashtags to track for such events, since the hashtags may not be used until after the event has occurred. Furthermore, *a priori* unknown events are not well localized in time, so it is hard to even determine when a critical transition happens. However, one may pose the simpler question of whether the scaling laws for the variance and autocorrelation warning signs can be detected in the time period preceding an *a priori known event*. This is a test of whether it may be possible to extract scaling laws from large-scale social networks near critical transitions at all. Since our data set encompasses fall/winter 2009 we chose the three special events:

(A) `#halloween`: Halloween occurred on Saturday, 31 October 2009

(B) `#thanksgiving`: Thanksgiving Day occurred on Thursday, 26 November 2009

(C) `#xmas`: Christmas Day occurred on Friday, 25 December 2009

Note that these events are well defined in the sense that the event dates are fixed and the hashtags are directly associated to the events since they do not commonly occur in different contexts. The first row of Fig. 1 shows the time series of frequency for each hashtag. As expected, the overall hashtag frequency count starts to increase weeks before the events (A)–(C). Note that we do know *a priori* that a drastic change is going to occur at the precise event date. However, if we would terminate the time

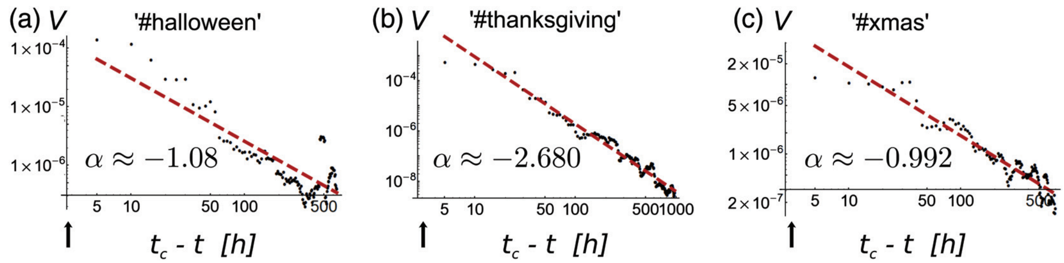

FIG. 2. Scaling exponents α from (1) extracted from the time series data using least squares fit (dashed red line) for the three hashtags from Fig. 1. Data are shown on a log–log scale with reverse time on the horizontal axis, and events are marked with arrows.

series several days or weeks before the event and did not know about (A)–(C) then we would have to take into account the possibility that a drastic spike never occurs and the word frequency just decreases slowly.

From this last perspective, i.e., by using only past data and information, we calculated the variance and lag-1 autocorrelation by considering sliding averages over a period of 50 h and time steps of 5 h. The results are shown in second and third rows of Fig. 1 respectively. In all three cases, we see a clear increase in the variance. The autocorrelation also increases before the events (A)–(C), as shown in row 3 of Fig. 1. Again, the warning sign is clearly visible for all three cases.

Furthermore, mathematical theory of critical transitions indicates that there is a regime in which we may expect to find a variance scaling law of the form:

$$\mathrm{Var}(t) = A(t_\mathrm{c} - t)^\alpha \quad \text{as } t \to t_\mathrm{c}, \tag{1}$$

where t_c denotes the time of the spiking event and where $A > 0$, $\alpha < 0$ are two constants. More details on the theory behind the asymptotic scaling law (1) are provided in Appendix A. However, we mention here that there are generically three different regimes for stochastic scaling laws: (R1) far away from the bifurcation point, (R2) a large-in-time regime approaching the bifurcation point and (R3) a small-in-time regime near the bifurcation. Theory predicts that a scaling law (1) should hold in regime (R2) and that the key measure is the exponent $\alpha < 0$ as it can potentially distinguish between different dynamical transitions. We performed a least-squares fit to determine α in (R2); see Fig. 2. We find that there is indeed a distinction between (R1), (R2) and (R3). Cases (A) and (C) show an exponent very close to $\alpha = -1$ which is an exponent generically expected near several types of bifurcation points such as Hopf, transcritical and pitchfork bifurcations. However, (B) shows a substantially smaller exponent than predicted by theoretical considerations. There could be several potential explanations for this non-standard exponent; see Appendix A. Further careful analysis is left for future work to determine whether the scaling law (1) is truly followed by the data [34,35]. Nevertheless, in all three cases we find a negative exponent α. This behavior is typical for the regime type (R2).

As the last aspect of our work, we also analysed the Twitter time series data to find examples of other events that are, e.g., unpredictable, exogenously induced, periodically driven, with short periods of warning signs or strongly noise-driven. Several alternative examples of such complementary behaviors are presented in Appendix C. These time series illustrate that there is a variety of potentially novel dynamical behaviors in large-scale social networks near large spikes that deserve to be investigated in their own right.

3. Discussion

We have demonstrated that clear warning signs exist for *a priori* known events in large-scale social media data. Our results leave open the possibility that warning signs could also exist for *a priori unknown* events. Investigating this problem further still remains a largely unexplored research topic. In particular, it is important to check various conjectures on the influence of new social media on major societal events very carefully.

Even very elementary questions are open in the context of social network data and warning signs. For instance, what type of time series data should be used? In this study, we analysed the content of messages posted by users, but one could equally well think of analysing the dynamics of connections made between users, or various other network activities. Even when considering the content of messages, options other than the topic or hashtag considered here are conceivable, such as particular words, word combinations or even entire phrases. Further important questions include:

- How do we define when a critical transition occurs in the data for an *a priori* unknown event?
- For *a priori* unknown events, is there a possibility to identify hashtags or other aspects of the message which allow us to determine the best warning sign?
- Can we link warning signs in social networks to *a priori* unknown critical transitions outside a social network?
- Which models of social networks can re-produce critical transitions observed in data?

In summary, our contribution is to provide a proof-of-principle. In particular, we have found some evidence for stochastic scaling laws, which are characteristic signs of critical transitions, in a large social media data set preceding large increases of activity. This suggests that the underlying social network may undergo a qualitative change between two different dynamic regimes. We explicitly emphasize that we neither claim to link these signs to any particular underlying mathematical network model nor can we prove for which types of network events such warning signs may exist; however, we have shown that there exist signatures that do resemble much the structures of warning signs observed in other fields, such as ecology, thus opening new perspectives on research on social network dynamics.

It is important to develop additional mathematical tools to answer further research questions. Due to their present lack, we have chosen examples of events that are sufficiently isolated and where influences, that we cannot quantify, are eliminated. Warning signs in social network data are still a largely unexplored topic. Indeed, it would be very desirable to have quantitative results explaining (i) the theoretical conditions under which we can expect drastic events in social networks to be predictable or unpredictable and (ii) which kind of practical demands must be met for the data. We expect that our work will stimulate further research, with ample opportunities to uncover results on social network dynamics with far-reaching ramifications.

Funding

This research was supported by the Austrian Academy of Sciences (ÖAW) via an APART fellowship and by the European Commission (EC/REA) via a Marie-Curie International Re-integration Grant (to C.K.). E.M.'s contribution is part of the Dynamical Systems Interdisciplinary Network, University of Copenhagen. This research was also supported by the Army Research Laboratory and accomplished under Cooperative Agreement Number W911NF-09-2-0053 and DARPA BAA-11-64, Social Media in Strategic Communication (to D.M.R.). We thank Jure Leskovec for providing data. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

1. Cohen, R. (2011) *Facebook and Arab dignity*. New York Times. Retrieved from http://www.nytimes.com/2011/01/25/opinion/25iht-edcohen25.html on 4 November 2013.
2. Webster, S. (2011) Has social media revolutionized revolutions? *World News* **87**. Retrieved from http://www.jcunews.com/2011/02/16/has-social-media-revolutionized-revolutions/ on 4 November 2013.
3. Hudson, L. & Flannes, M. (2011) *The Arab Spring: Anatomy of a tipping point*. Aljazeera. Retrieved from http://www.aljazeera.com/indepth/opinion/2011/08/201183081433165611.html on 4 November 2013.
4. Lim, M. (2012) Clicks, cabs, and coffee houses: social media and oppositional movements in Egypt, 2004–2011. *J. Commun.* **62**, 213–248.
5. Tufekci, Z. & Wilson, C. (2012) Social media and the decision to participate in political protest: observations from Tahrir Square. *J. Commun.* **62**, 363–379.
6. Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I. & Boyd, D. (2011) The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *Int. J. Commun.* **5**, 1375–1405.
7. Lim, M. (2011) Tahrir Square was a foreseeable surprise: tracing the history of Egyptian online activism. *Slate*. Retrieved from http://www.slate.com/articles/technology/future_tense/2011/07/tahrir_square_was_a_foreseeable_surprise.html on 4 November 2013.
8. Bimber, B., Flanagin, A. J. & Stohl, C. (2005) Reconceptualizing collective action in the contemporary media environment. *Commun. Theory* **15**, 365–388.
9. Agarwal, N., Lim, M. & Wigand, R. (2012) Raising and rising voices in social media - a novel methodological approach in studying cyber-collective movements. *Bus. Inform. Syst. Eng.* **4**, 113–126.
10. Scheffer, M., Bascompte, J., Brock, W. A., Brovkhin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M. & Sugihara, G. (2009) Early-warning signals for critical transitions. *Nature* **461**, 53–59.
11. Kuehn, C. (2011) A mathematical framework for critical transitions: bifurcations, fast-slow systems and stochastic dynamics. *Physica D* **240**, 1020–1035.
12. Dai, L., Vorselen, D., Korolev, K. S. & Gore, J. (2012) Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* **336**, 1175–1177.
13. Drake, J. M. & Griffen, B. D. (2010) Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459.
14. Veraart, A. J., Faassen, E. J., Dakos, V., van Nes, E. H., Lurling, M. & Scheffer, M. (2012) Recovery rates reflect distance to a tipping point in a living system. *Nature* **481**, 357–359.
15. Wang, R., Dearing, J. A., Langdon, P. G., Zhang, E., Yang, X., Dakos, V. & Scheffer, M. (2012) Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature* **492**, 419–422.
16. Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W. A., Cline, T., Coloso, J., Hodgson, J. R., Kitchell, J. F., Seekell, D. A., Smith, L. & Weidel, B. (2011) Early warning signs of regime shifts: a whole-ecosystem experiment. *Science* **332**, 1079–1082.
17. Kuehn, C. (2013) A mathematical framework for critical transitions: normal forms, variance and applications. *J. Nonlinear Sci.* **23**, 457–510.
18. Ashwin, P., Wieczorek, S., Vitolo, R. and Cox, P. (2012) Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Phil. Trans. R. Soc. A* **370**, 1166–1184.
19. Dakos, V., van Nes, E. H., Donangelo, R., Fort, H. & Scheffer, M. (2009) Spatial correlation as leading indicator of catastropic shifts. *Theoret. Ecol.* **3**, 163–174.

20. Lade, S. J. & Gross, T. (2012) Early warning signals for critical transitions: a generalized modeling approach. *PLoS Comput. Biol.* **8**, e1002360–6.

21. Lenton, T. M. (2011) Early warning of climate tipping points. *Nat. Clim. Change* **1**, 201–209.

22. Thompson, J. M. T. & Sieber, J. (2011) Climate tipping as a noisy bifurcation: a predictive technique. *IMA J. Appl. Math.* **76**, 27–46.

23. Cimatoribus, A. A., Drijfhout, S. S., Livina, V. & van der Schrier, G. (2012) Dansgaard-Oeschger events: tipping points in the climate system. *Clim. Past Discuss.* **8**, 4269–4294.

24. Venegas, J. G., Winkler, T., Musch, G., Vidal Melo, M. F., Layfield, D., Tgavalekos, N., Fischman, A. J., Callahan, R. J., Bellani, G. & Harris, R. S. (2005) Self-organized patchiness in asthma as a prelude to catastrophic shifts. *Nature* **434**, 777–782.

25. Meisel, C. & Kuehn, C. (2012) On spatial and temporal multilevel dynamics and scaling effects in epileptic seizures. *PLoS ONE* **7**, e30371.

26. Chertkov, M., Backhaus, S., Turzisyn, K., Chernyak, V. & Lebedev, V. (2011) *Voltage collapse and ODE approach to power flows: analysis of a feeder line with static disorder in consumption/production*. *arXiv:1106.5003v1*, pp. 1–8.

27. O'Regan, S. M. & Drake, J. M. (2013) Theory of early warning signals of disease emergence and leading indicators of elimination. *Theoret. Ecol.* **6**, 333–357.

28. Carpenter, S. R. & Brock, W. A. (2006) Rising variance: a leading indicator of ecological transition. *Ecol. Lett.* **9**, 311–318.

29. Ditlevsen, P. D. & Johnsen, S. J. (2010) Tipping points: early warning and wishful thinking. *Geophys. Res. Lett.* **37**, 19703.

30. Gladwell, M. (2002) *The Tipping Point: How Little Things Can Make a Big Difference*. Back Bay Books.

31. Lamberson, P. J. & Page, S. E. (2012) Tipping points. *Int. Quart. J. Polit. Sci.* **7**, 175–208.

32. Slater, D. Early warning signals of tipping-points in blog posts. Preprint: 2012. Retrieved from http://www.mitre.org/work/tech_papers/2012/12_4711/12_4711.pdf on 4 November 2013.

33. Yang, J. & Leskovec, J. (2011) Patterns of temporal variation in online media. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, pp. 177–186.

34. Clauset, A., Shalizi, C. R. & Newman, M. E. J. (2009) Power–law distributions in empirical data. *SIAM Rev.* **51**, 661–703.

35. Stumpf, M. P. H. & Porter, M. (2012) Critical truths about power laws. *Science* **335**, 665–6.

36. Kuznetsov, Y. A. (2004) *Elements of Applied Bifurcation Theory*, 3rd edn. New York: Springer.

37. Guckenheimer, J. & Holmes, P. (1983) *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York, NY: Springer.

38. Kuehn, C. (2009) Scaling of saddle-node bifurcations: degeneracies and rapid quantitative changes. *J. Phys. A* **42**, 045101.

39. Berglund, N. & Gentz, B. (2006) *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*. Berlin: Springer.

Appendix A. Critical transitions/tipping points

In this section we provide some technical background on critical transitions (or tipping points) to support our main theme of studying variance and autocorrelation near jumps in time series of complex networks. Consider a general n-dimensional system of ordinary differential equations (ODEs) given by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, \tilde{\mu}), \quad (x, \tilde{\mu}) \in \mathbb{R}^m \times \mathbb{R}^p, \tag{A.1}$$

where we may view $\tilde{\mu}$ as the parameters describing the network properties and x as all the relevant variables to track the state of the network; the map f is assumed to be sufficiently smooth. Suppose there exists a branch of equilibria $x^* = x^*(\tilde{\mu})$, i.e., steady states with $f(x^*(\tilde{\mu}), \tilde{\mu}) = 0$. Suppose the equilibria are stable for a certain range of parameter values and lose stability for some $\tilde{\mu} = \tilde{\mu}^*$ at which the state of the system (i.e. the network) changes drastically. In this case it can be proved rigorously that the only two situations which imply stability loss and are stable to arbitrary perturbations within the class of sufficiently smooth vector fields upon variation of a *single* parameter are the so-called fold and Hopf bifurcations (see the description of such (codimension 1) bifurcations in [36, Chapter 3] as well as [37, Section 3.4]). Due to this theory we may start from the low-dimensional cases $(m, p) = (1, 1)$ for the fold and $(m, p) = (2, 1)$ for the Hopf bifurcation. The fold is always a tipping point while the Hopf bifurcation only induces tipping to a distant attractor in the subcritical case [11, Proposition 2.5–2.6]. We remark that one can show that it is a generic property [17, Section 7.4] that the warning signs we are going to discuss appear in all variables $x \in \mathbb{R}^m$ upon sufficient coupling i.e. before reducing to the minimal dimensions [37, Section 3.2].

For the fold bifurcation one may prove [36, Section 3.2] that the system near the transition, where a stable and unstable steady state collide and annihilate, may be reduced to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu - x^2 + \mathcal{O}(\cdots), \quad (x, \mu) \in \mathbb{R} \times \mathbb{R}, \tag{A.2}$$

where $\mathcal{O}(\cdots)$ denotes higher-order terms which we shall drop from now on and, for concreteness, we may think of x as the word frequency. The two steady states are $x^\pm = \pm\sqrt{\mu}$ for $\mu > 0$ and x^+ is stable while x^- is unstable. The fold bifurcation occurs for $\mu = 0$. Abstractly, one may view the stable steady state x^+ as the representation of the word frequency in the network before it approaches the drastic jump upon parameter variation of μ (which we could just take as progression of time for our case). Linearizing around x^+ yields the variational equation

$$\frac{\mathrm{d}X}{\mathrm{d}t} = (D_x f)(x^+)X = -2\sqrt{\mu}X \quad \Rightarrow X(t) = X(0)\,\mathrm{e}^{-2\sqrt{\mu}t}. \tag{A.3}$$

From (A.3) it follows that the local linear exponential stability of x^+ decreases as $\mu \to 0$; this is just the classical slowing-down phenomenon (or intermittency) [37, pp. 343–346]. If one perturbs (A.2) by an additive noise process $\xi(t)$ which is delta-correlated white noise, i.e., $\mathbb{E}[\xi(t)] = 0$, $\mathbb{E}[\xi(t_1)\xi(t_2)] = \delta(t_1 - t_2)$, and where $\sigma > 0$ is a parameter controlling the noise level, then this yields the stochastic differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu - x^2 + \sigma\xi(t), \tag{A.4}$$

which is to be interpreted in the sense of the Itô-calculus. Under certain assumptions on the parameter variation of μ and the sufficiently small noise level σ [17, p. 479] one may prove that there is a region before the fold where the time series fluctuating near the deterministic steady state exhibits a stochastic scaling law for the variance

$$\mathrm{Var}(x(t)) = \sigma^2\mathcal{O}\left(\frac{1}{\sqrt{\mu}}\right) + \text{higher-order error terms} \sim \mathcal{O}\left(\frac{1}{\sqrt{\mu}}\right) \tag{A.5}$$

as $\mu \to 0$ from above and $\sigma > 0$ remains fixed. A related result, obtained via a fluctuation–dissipation relation, is that the autocorrelation increases as $\mu \to 0$ from above; see [10,29] for a discussion of this aspect. The argument for the case of the Hopf bifurcation follows in an analogous manner and the resulting scaling law turns out to be [17, Theorem 5.2]

$$\mathrm{Var}(x(t)) = \sigma^2 \mathcal{O}\left(\frac{1}{\mu}\right) + \text{higher-order error terms} \sim \mathcal{O}\left(\frac{1}{\mu}\right). \tag{A.6}$$

Hence, this motivates to test for an increase in variance and autocorrelation in the time series data. Furthermore, replacing the parameter μ by a time-dependent drift, or just by time t for simplicity, we expect to detect asymptotic scaling laws of the form:

$$\mathrm{Var}(x(t)) \sim A(t_c - t)^{\alpha}, \quad t \to t_c, \tag{A.7}$$

where t_c is the time when the critical transition occurs. The two generic scaling exponents are $\alpha = -1$ and $-\frac{1}{2}$. However, if we additionally drop the assumption of additive noise and instead assume multiplicative noise, which depends on the distance to the critical transition point, then the two generic 1-parameter scaling exponents can decrease or increase as demonstrated in [17, Section 7.5]. This is one scenario which could potentially account for the increased exponent observed in Fig. 2(b).

Although the abstract mathematical analysis can be justified rigorously, it has to rely on a number of assumptions to achieve these results, including e.g. smoothness of the ODEs [38], coupling of the normal form variable to the measured variable [17, Section 7.4], weak-noise regime in comparison to parameter drift [39, p. 96] and additive noise (as discussed above). The current state of abstract social network modelling does not suffice to check all these assumptions for a large realistic model. Interestingly, it has been shown in the context of classical coarse-grained epidemic models that these assumptions may hold in many cases; notably, epidemic models could potentially share many features with social networks, see [17, Section 7.2] and [27]. In any case, variance and autocorrelation are relatively easy to measure using a sliding window approach outlined in the previous section.

In conclusion, this provides the mathematical motivation for our basic hypothesis that we would like to test for in this study: Are there *any* signs of increasing variance and autocorrelation before spikes/jumps in network activity?

Appendix B. Time-series analysis

We briefly describe the time-series analysis. The basic time unit we use is the number of hours elapsed since the starting point of the time series. Denote the word frequency during one hour x_t. For example, for `#xmas`, we have

$$x_t = \frac{\text{number of occurrence of \texttt{\#xmas} during } [t-1, t]}{\text{tweet volume during } [t-1, t]}$$

which yields a time series $\{x_1, x_2, \ldots, x_t, \ldots\}$. To compute the warning signs we fix a window size n. At a time point $t_k > n + l$ for some fixed lag $l \geqslant 1$ we consider the vector

$$\mathbf{x_k} := (x_{t_k-(n-1)}, x_{t_k-(n-2)}, \ldots, x_{t_k}).$$

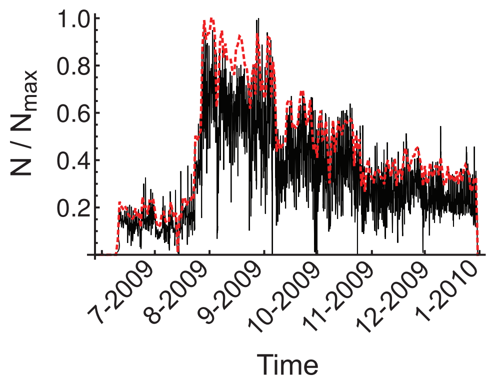

FIG. 3. Relative tweet volume $N/N_{\max}$ per hour (black) and per day (red) during the period from June 1 2009 to December 31 2009.

Denote the vector obtained by removing the linear trend from $\mathbf{x_k}$ by $\mathbf{y_k}$. Then we may just compute the mean μ_k, variance V_k and lag-l autocorrelation $R_{k,l}$ of $\mathbf{y_k}$ by the standard formulae

$$\mu_k = \frac{1}{n}\sum_{j=1}^{n}(y_k)_j, \qquad V_k = \frac{1}{n}\sum_{j=1}^{n}((y_k)_j - \mu_k)^2,$$

$$R_{k,l} = \frac{1}{(n-l)V_k}\sum_{j=1}^{n-l}(y_j - \mu_k)(y_{j+l} - \mu_k).$$

This yields the required time series for the warning signs given by V_k and $R_{k,l}$. Note that we also choose a time step s for the spacing in the index k which yields that

$$\mathbf{V} = (\ldots, V_k, V_{k+s}, V_{k+2s}, \ldots)$$

and similarly for $\mathbf{R_l}$. To understand why the variance and autocorrelation are expected to increase generically near certain types of critical transitions we refer the reader to [10,11,17] where also the specific scaling laws are discussed.

The relative tweet volume N/N_m per hour and day are shown in Fig. 3; the maximum recorded volume is $N_m = 5,106,720$ per hour or $N_m = 296,256$ per day, respectively. The average tweet activity is 2.2165×10^6 per day or 92767.7 per hour, with a standard deviation of 1.27554×10^6 per day or 59375.9 per hour, respectively.

Appendix C. Further examples of time series

To illustrate various other phenomena that can be found in large-scale social media, time series, we discuss a few additional hashtag time series, shown in Figs 4 and 5. Another example exhibiting warning signs for an *a priori* known event is the time series for `#spring` in Fig. 4(a). However, note that the event of ‘spring’ is not confined to a single particular date and therefore, we observe a broad regime of large activity instead of one distinct spike. Figure 4(b) shows the time series for `#trick`, which has a very large peak at Halloween (marked by a dashed red line), probably due to its relation to the phrase ‘trick-or-treat’; but the variance and autocorrelation begin to increase only very close to the spike in

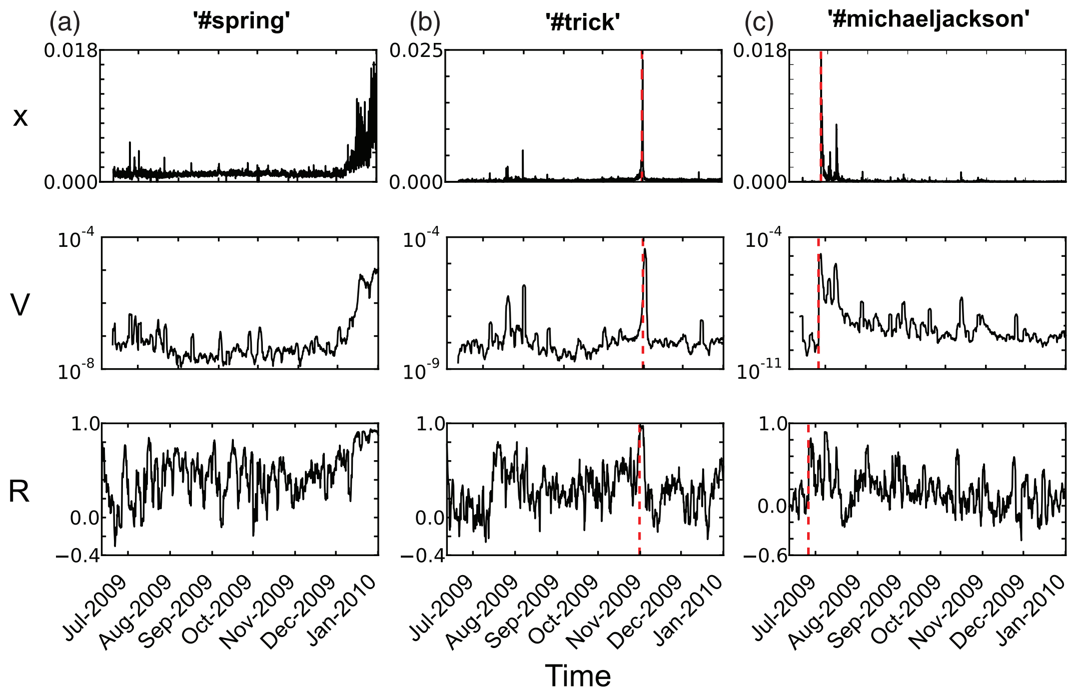

FIG. 4. Further examples of time series with large-spike events. Columns display from top to bottom normalized word frequency x, variance V and lag-1 autocorrelation R calculated from the time series of word frequency. (a) #spring is another example of a time series with clear early warning signs; in fact, the event ('spring') is again *a priori* known, but by its own nature, it is not an as well-defined singular event as the examples shown in Fig. 1. (b) #trick shows a larger spike which is difficult to predict and thus may evade the time-series analysis discussed here. (c) The sudden death (date marked) of pop-star Michael Jackson illustrates an example of a sudden time series spike that cannot be anticipated due to its inherently exogenous nature.

the time series. This illustrates how indirect indicators of an event could be substantially worse predictors than the hashtags for the event themselves. Figure 4(c) shows the hashtag `#michaeljackson` exhibiting an extremely large spike at the day of the unexpected death of Michael Jackson, but—obviously—without any warning signs preceding the event. The cause of this event is clearly exogenous in nature (in terms of social network dynamics) and thus, it is impossible to detect (or anticipate) early warning signs in these time series data.

Figure 5(a) shows the time series for the hashtag `#windows_7`, related to the release event (red dashed line) of the software package Windows 7. There is increasing variance as well as autocorrelation before the (*a priori* known) event. However, the warning signs are not as clear-cut as in the examples shown in Fig. 1. In fact, the prediction time window seems rather short. Thus, this example represents a case which lies in between the realm of predictable and unpredictable events, at least when the detection of warning signs is based upon the variance and autocorrelation analysis. Figure 5 shows the time series for the hashtag `#flu`, which is very noisy. Although there are several spikes in activity, these spikes appear clearly unpredictable as—from a modelling point of view—seem to be noise-induced events. Finally, Fig. 5 shows a periodically driven time series for `#friday`. Warning signs could potentially be possible to identify during each week before the spikes, but we do not examine this case in further detail since, to the best of our knowledge, there is no theoretical background for warning signs for systems with short rapid periodic forcing.

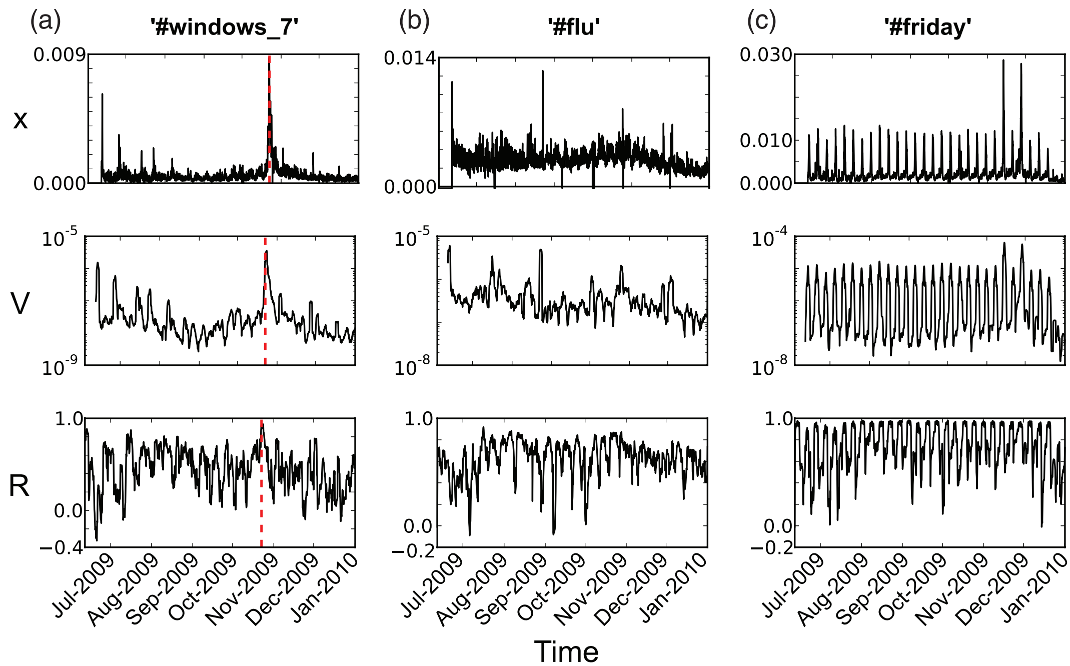

FIG. 5. Further examples of time series with large-spike events. Columns display from top to bottom normalized word frequency x, variance V and lag-1 autocorrelation R calculated from the time series of word frequency. (a) The red vertical dashed line indicates the *a priori* known date of the release event of the software package Windows 7. (b) The time series for #flu displays high noise levels where occasional larger spikes cannot be predicted. (c) The time series for #friday exhibits clear periodic driving for self-evident reasons.