Evolution of the digital society reveals balance between viral and mass media influence

Kaj Kolja Kleineberg*, Marián Boguñá

Dept. de Física Fonamental – Universitat de Barcelona – *kkl@ffn.ub.edu

Abstract

Online social networks (OSNs) enable researchers to study the social universe at a previously unattainable scale. The worldwide impact and the necessity to sustain their rapid growth emphasize the importance to unravel the laws governing their evolution. Empirical results show that, unlike many real-world growing networked systems, OSNs follows an intricate path that includes a dynamical percolation transition. At the light of these results, we present a quantitative two-parameter model which reproduces the entire topological evolution of a quasi-isolated OSN with unprecedented precision from the birth of the network. This allows us to precisely gauge the fundamental macroscopic and microscopic mechanisms involved. Our findings suggest that the coupling between the real pre-existing underlying social structure, a viral spreading mechanism, and mass media influence govern the evolution of OSNs. The empirical validation of our model, on a macroscopic scale, reveals that virality is four to five times stronger than mass media influence and, on a microscopic scale, individuals have a higher subscription probability if invited by weaker social contacts, in agreement with the "strength of weak ties" paradigm.

Case study

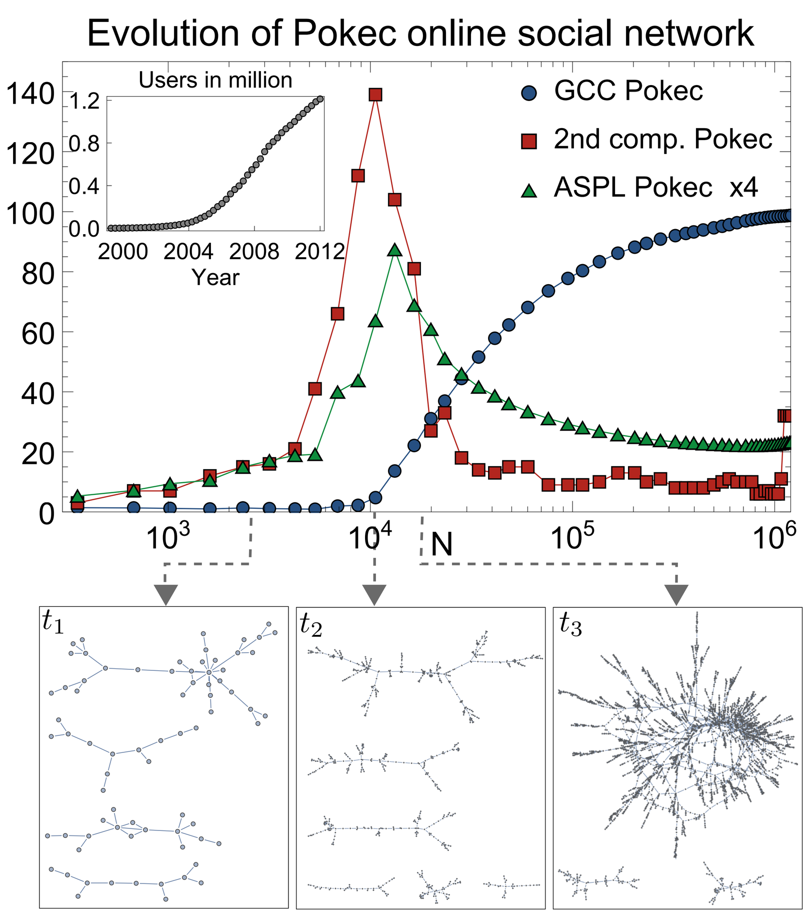

Fig. 1: Evolution of Pokec network

We conducted a case study about Pokec, a large slovakian online social network with 1.6 users and directed links. Nevertheless, not all directed links correspond to a real social tie. Thus, we discard all non-bidirectional links from the original graph and treat those left as undirected edges. The resulting filtered network is composed of 1.2 million users. The Pokec network has the following unique features:

1. Isolated (in Slovak language)
2. Large population coverage
3. Successful (ongoing growth)

We reconstruct the evolution of the network from its birth. We observe a dynamical phase transition between a phase which consists of small disconnected clusters and a percolated phase where a macroscopic fraction of the network is connected. See Fig 1.

Basic model

We propose a two-layer multiplex model where the bottom layer corresponds to the pre-existing social structure and the upper layer represents the OSN (see Fig. 2). We use the final snapshot of the empirical network as the underlying network because it is a good proxy for the real social structure due to its large population coverage. Nodes can be in the following three states: active, passive, or susceptible. Susceptible nodes are only present in the underlying layer, whereas active and passive nodes are present in both layers. The model incorporates the following dynamical processes:

1. *Viral activation:* a susceptible node can be virally activated and added to the OSN by contact to an active neighbor in the traditional off-line network. This event happens at rate λ per each active link.
2. *Mass media effect:* each susceptible individual becomes active spontaneously at rate μ and is added to the OSN layer as a response to the visibility of the OSN.
3. *Deactivation:* active users become spontaneously passive at rate δ (= 1) and no longer trigger viral activations nor reactivate other passive nodes.
4. *Viral reactivation:* at rate λ an active user can reactivate a passive neighbor. The neighbor then becomes active and can trigger both viral activations and viral reactivations.

We match the system size at the critical point for the model and the data (see Fig. 3), which yields a linear relation between the parameters λ and μ. In Fig. 4 the model results are shown. The model performs very well for the considered metrics.

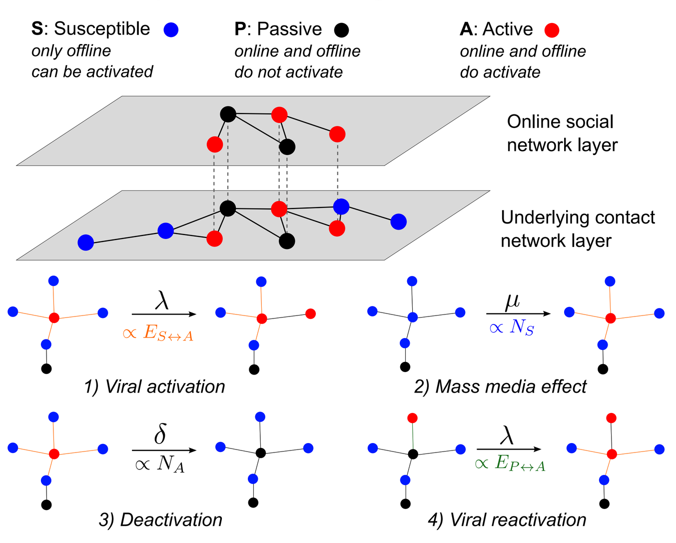

Fig. 2: Two-layer model design

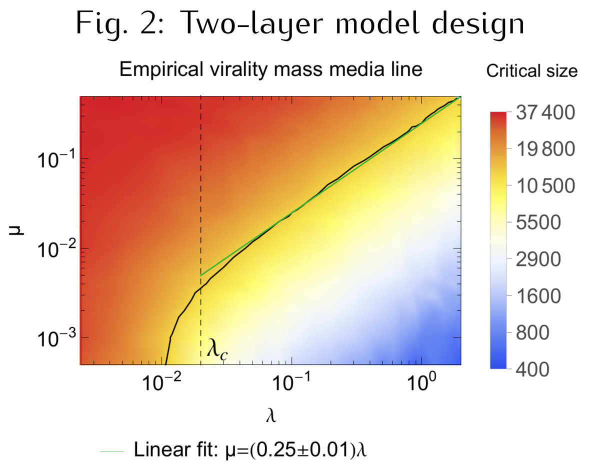

Fig. 3: Balance of parameters λ and μ

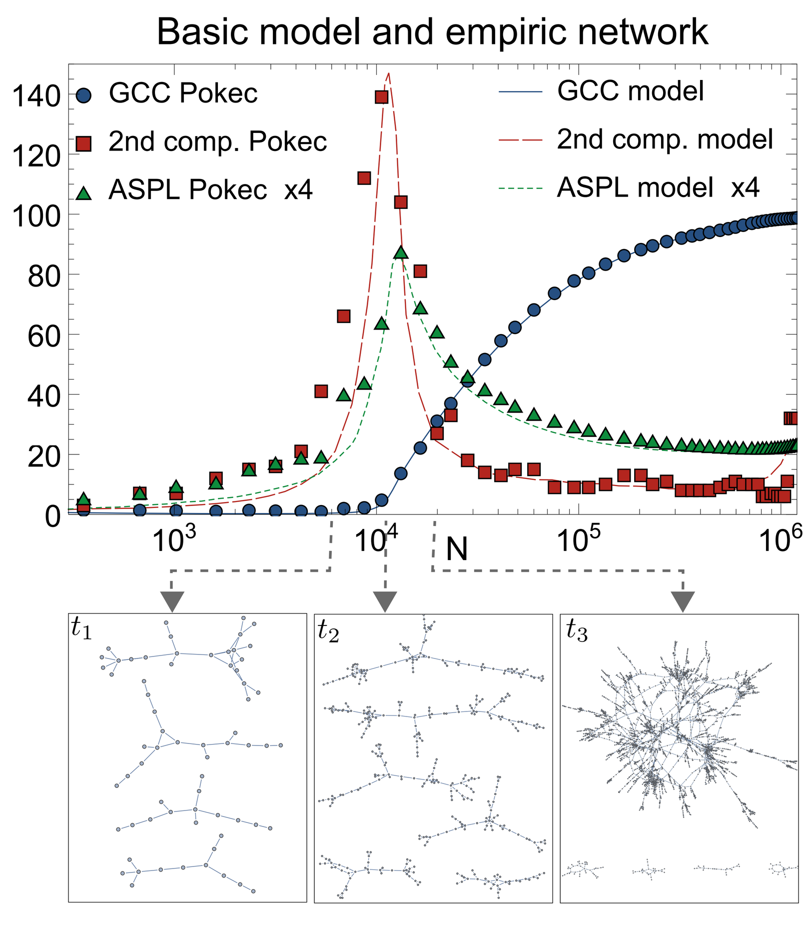

Fig. 4: Model results and comparison with Pokec

Extended model

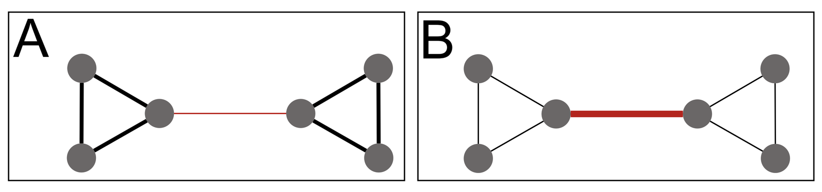
Fig. 5: Transmissibility as a function of tie strength. Left corresponds to $\eta > 0$, right to $\eta < 0$

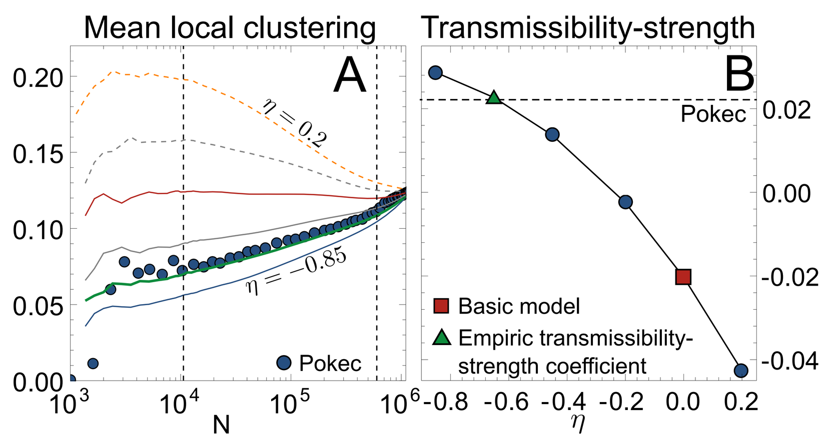

Fig. 6: Transmissibility-strength relation

The basic model works very well for the metrics considered so far. However, for local quantities the model fails to reproduce the observations. For instance, the local clustering coefficient in the model exhibits a decreasing trend whereas it increases for the empiric network (see Fig. 7D). We extend the model by taking into account the local neighborhood of each link (so far, we treated all links equally). To each link, we assign the viral transmissibility

$$\lambda_{ij} = \lambda \frac{s_{ij}^{\eta}}{\left\langle s_{ij}^{\eta} \right\rangle}$$

where $s_{ij} = m_{ij} + 1$ is the multiplicity of the edge plus 1. See Fig. 5 for an illustration.

The viral activation and reactivation along the link between i and j now occur at rate λ_{ij}. The exponent η controls the transmissiblity-strength relation and impacts the trend of the clustering coefficient. In Fig. 6A, we show the evolution of the clustering coefficient for different values of η. The best matching with the empirical data is obtained for $\eta = -0.65$ (Fig. 6B). Interestingly, it is a negative value which means individuals are more likely to join if invited by weaker social contacts. Our findings are in accordance with recent empirical observations about "structural diversity" and Granovetters theory "the strength of weak ties".

Further results and comparison

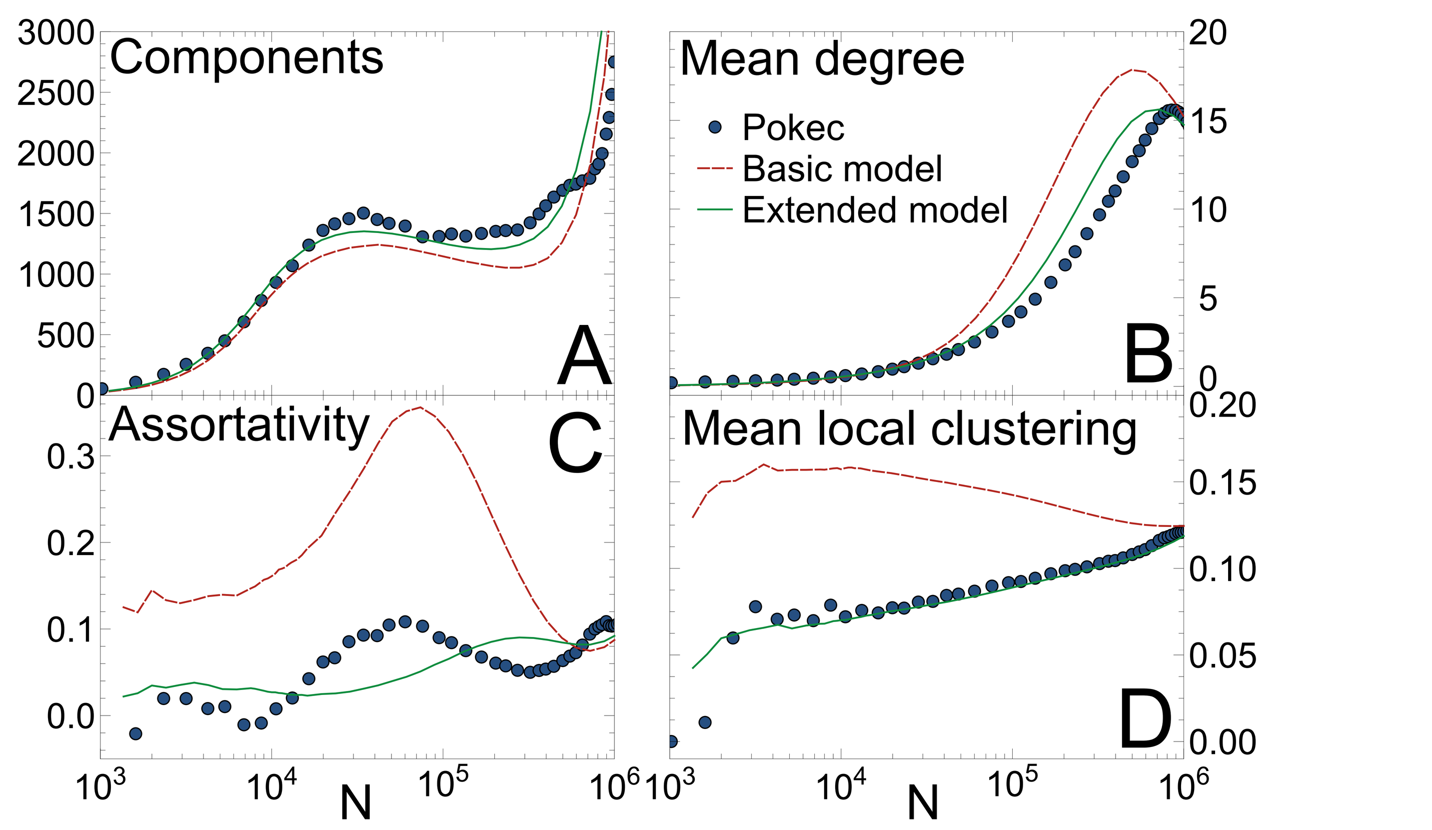

Fig. 7: Evolution of A: the number of components of size > 1, B: mean degree, C: assortativity coefficient and D: clustering coefficient for the basic and extended model compared with the Pokec network.

Summary and discussion

Comprehensive datasets on the evolution of OSNs offer us the opportunity to determine the principal mechanisms involved in social contagion and online activity of individuals. At this respect, the OSN Pokec, with its peculiar evolution and being almost isolated, is particularly appropriate. Interestingly, the evolution of Pokec's topology is characterized by a dynamical percolation transition, a rather peculiar behavior in real evolving networks. We have shown that this anomalous topological evolution can be explained very precisely on a quantitative level by a two-layer model, which accounts for the underlying real social structure, combined with two main mechanisms. First, a viral effect, responsible for the social contagion of new users and, second, a mass media effect, leading to random subscriptions of new users. Interestingly, the balance between these two mechanisms is what governs the topological growth of OSNs. In the particular case of the OSN Pokec, the quantification of this balance reveals that the viral effect is between four to five times stronger than the mass media effect. This can explain the proliferation of viral marketing campaigns, in detriment of traditional advertising. To our knowledge for the first time a model with only very few parameters yields quantitatively precise insights about the topological formation of OSNs.

Beyond the global behavior of our basic model, the social neighborhood of individuals has shown to be crucial to explain the evolution of local topological quantities in Pokec. We find that viral transmissibility is inversely proportional to the strength of social ties. This result is particularly interesting as it corroborates recent empirical findings concerning the role of "structural diversity" on social contagion processes by analyzing email invitations from Facebook users. However, our model allows us to identify and quantify this effect exclusively from –and hence its impact on– the topological evolution of the OSN. Alongside with Granovetter's conclusion about the importance of weak ties for individual success, our results give rise to the interpretation that OSNs evolve in a way to improve the possibilities for individual success. This might constitute an important reason for the huge popularity of OSNs.

Our findings here suggest interesting future research lines. Indeed, the particular OSN analyzed in this paper is a quasi-isolated system and, thus, allows us to gauge the fundamental mechanisms at play in the evolution of OSNs. However, in a general situation, an entire ecosystem of OSNs operate simultaneously, competing for the same users, which now become a scarce resource. The introduction of competition among OSNs in our model opens the possibility to develop an ecological theory of the digital world.

For further information, see Phys. Rev. X, 4:031046 (2014).