Supporting Information for Evolution of the Digital Society Reveals Balance between Viral and Mass Media Influence

Kaj-Kolja Kleineberg[1] and Marián Boguñá[1]

[1]*Departament de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain*

(Dated: September 10, 2014)

Keywords:

I. NULL MODEL

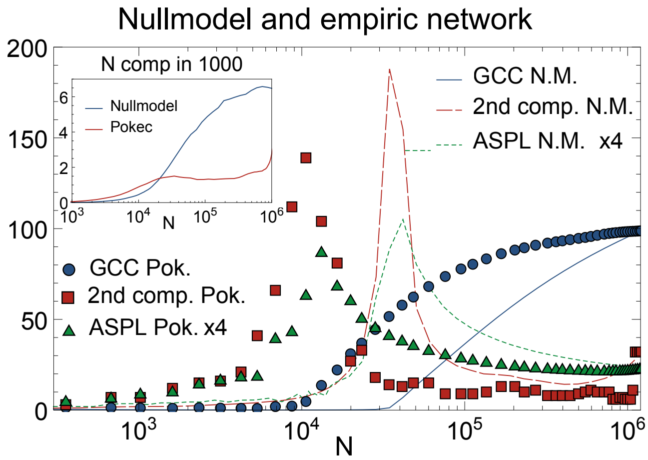

Figure 1: A null model with underlying Pokec network which consists of randomly adding nodes to the online social network. Points correspond to the empiric network and solid lines correspond to the null model.

In figure 1 we present the results from a null model for the role of the underlying empiric network. We again use the Pokec network as underlying and add nodes completely randomly to the network. Note, that this corresponds to our model for $\lambda = 0$ and arbitrary $\mu > 0$. However, the choice of μ then just fixes the model timescale, which we adjust implicitly by transforming physical time to the intrinsic network timescale given by the number of nodes. We observe, that the phase transition takes place at a larger network size. Note, that there is no more parameter to adjust. The number of components (see inset in fig. 1) also varies strongly between the empiric network and the presented null model.

We conclude, that the occurrence of the phase transition is included in the structure of the underlying network. Nevertheless, a null model with exclusively random subscriptions fails to reproduce the critical point of the phase transition as well as the evolution of the number of components.

II. SUSTAINED ACTIVITY THRESHOLD

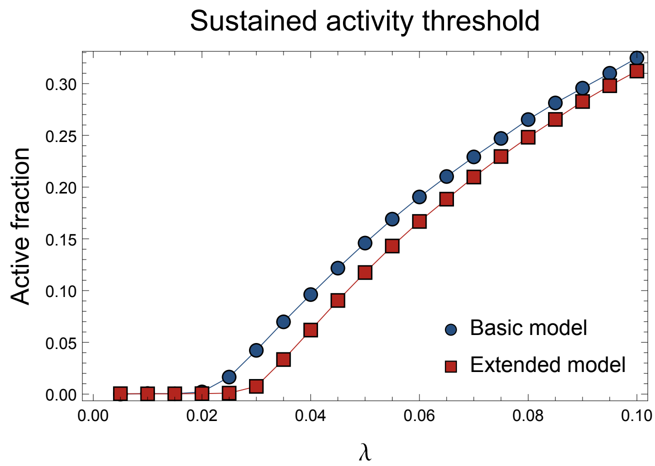

Figure 2: Sustained activity threshold $\lambda_c \approx 0.02$. Below this threshold, the activity of the network is not sustained and eventually the whole network will become passive.

From figure 2 we obtain the critical parameter λ_c for the outbreak of the SIS equivalent PAP dynamics within the online social network layer. We obtain

$$\lambda_c \approx 0.02 \tag{1}$$

which appears to be quite robust to the assignment of weights. Below this threshold, the whole network become passive. We suggest, that this corresponds to the practical disappearance of the network as observed in many empiric online social networking services.

III. PATHLENGTH AND DIAMETER

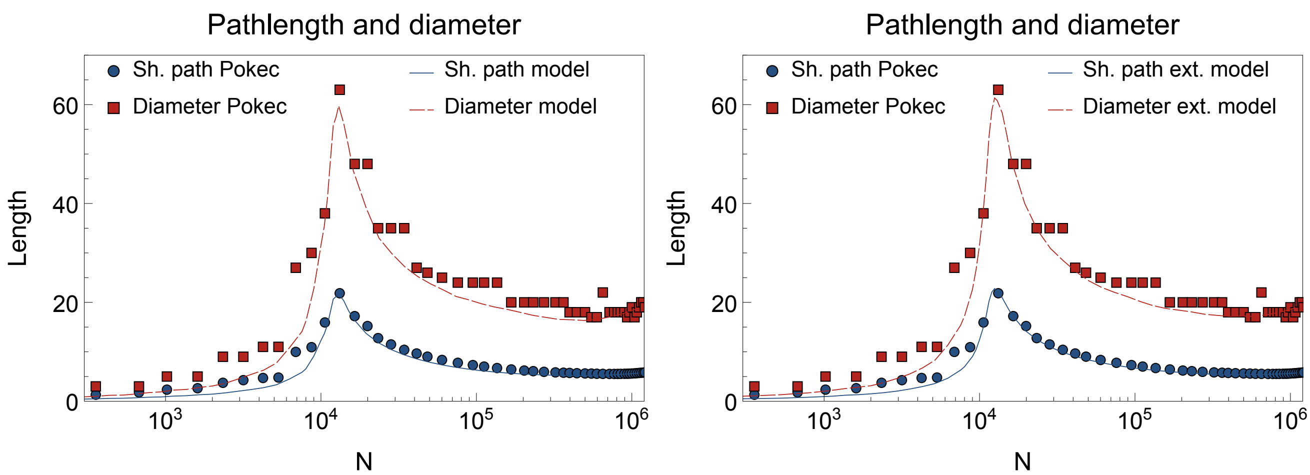

Figure 3: Pathlength and diameter for basic model (left) and extended model (right).

We observe the same behavior in the evolution of the average shortest path length and the network diameter (see figure 3). In the connected regime the pathlength and the diameter within the GCC decrease. In the disconnected regime, the average shortest path length and the diameter increase and reach their maximum at the critical point. The behavior of the extended model is equivalent.

IV. AVERAGE FINITE CLUSTER SIZE

Alternatively to the size of the second largest component one can consider the average finite cluster size, which is the average size of disconnected components without the largest one. The average finite cluster size also exhibits a peak at the critical point. Here, we show the average finite cluster size for all components with size $N > 1$.

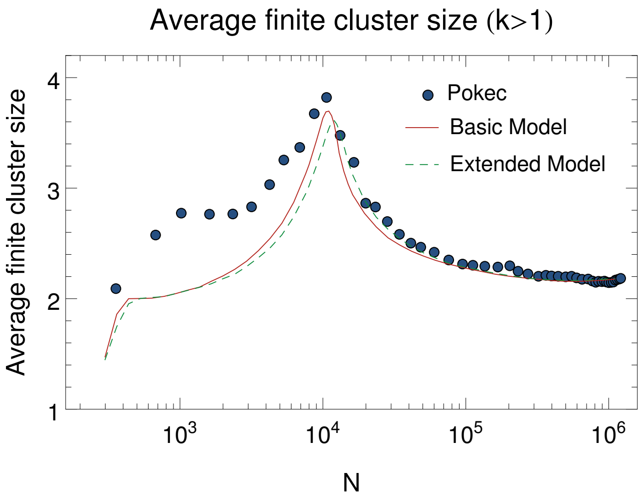

Figure 4: The evolution of the average finite cluster size for the Pokec network, the basic model, and the extended model for the same parameters as used in the paper.

V. EXPLICIT RESULTS FOR EXTENDED MODEL

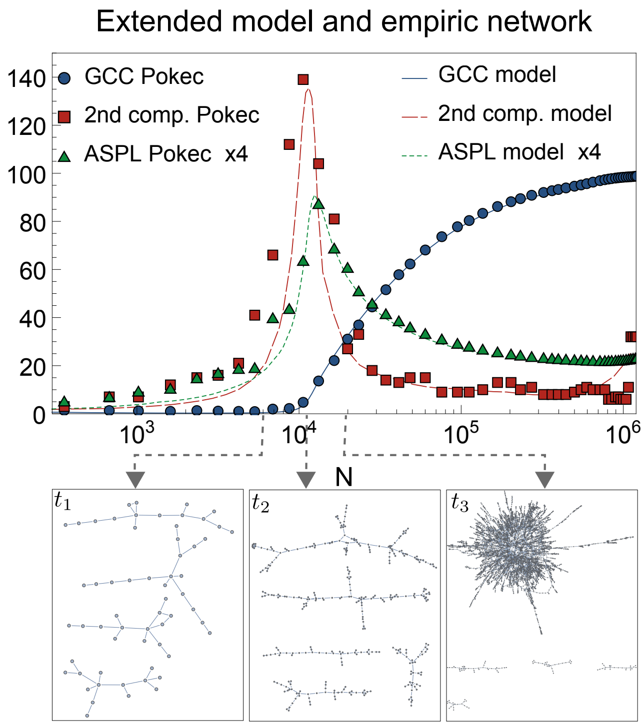

Figure 5: GCC (blue), size of second largest component (red), and average shortest path length (green) for the extended model for the parameters $\eta = -0.65$, $\lambda = 0.03$, and $\mu = 0.006$.

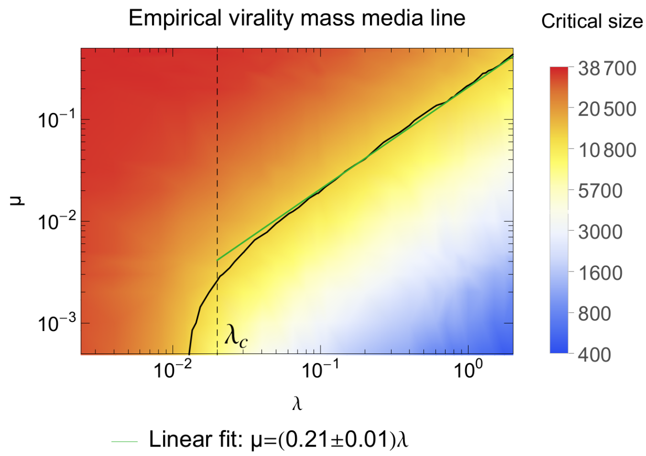

Figure 6: Matching of the critical point for the extended model.

VI. DEGREE DISTRIBUTION OF POKEC OSN

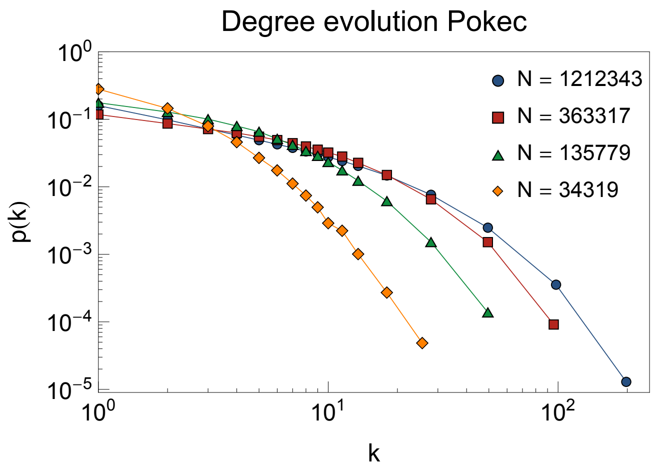

Figure 7: Degree distribution of Pokec OSN for different times.

VII. DISTRIBUTION OF COMPONENT SIZES

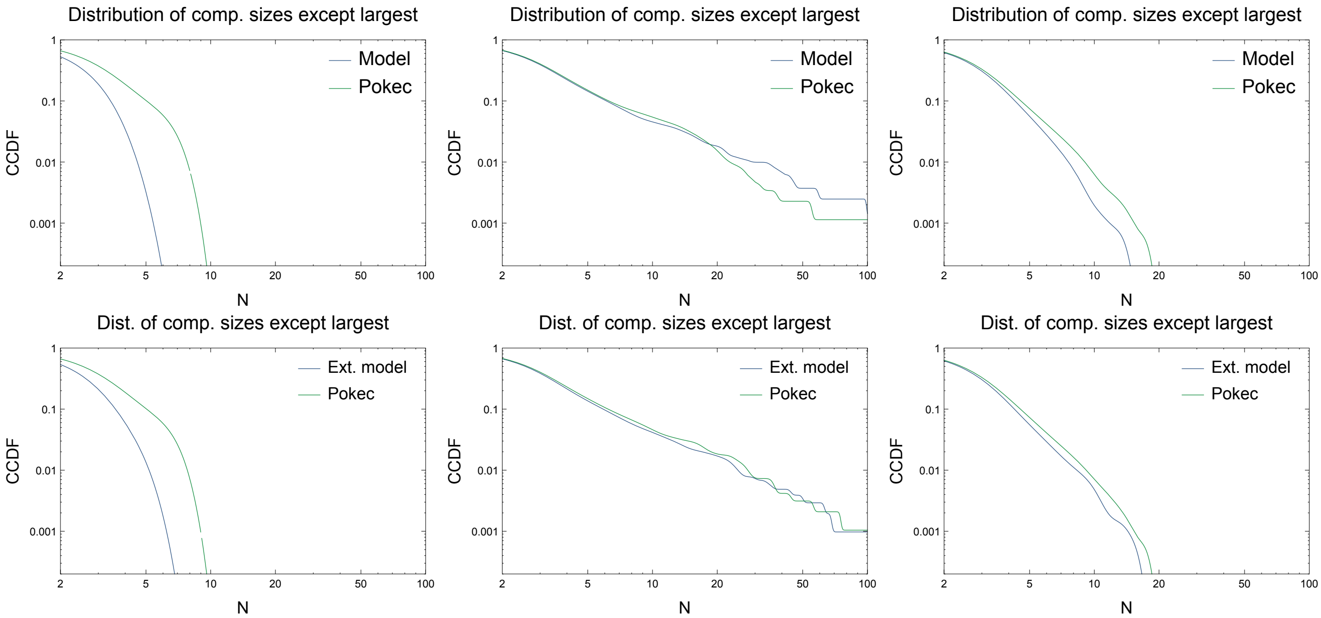

Figure 8: Distribution of component sizes except largest component for the basic model (top row) and for the extended model (bottom row). Network sizes are the following. **Left column:** $N = 1000$, **center column:** $N = 10000$, **right column:** $N = 29000$. The center column shows the distribution of sizes near the critical point. One sees that the distribution follows a power-law which is expected at the critical point of a phase transition.

VIII. DEMOGRAPHICS ANALYSIS

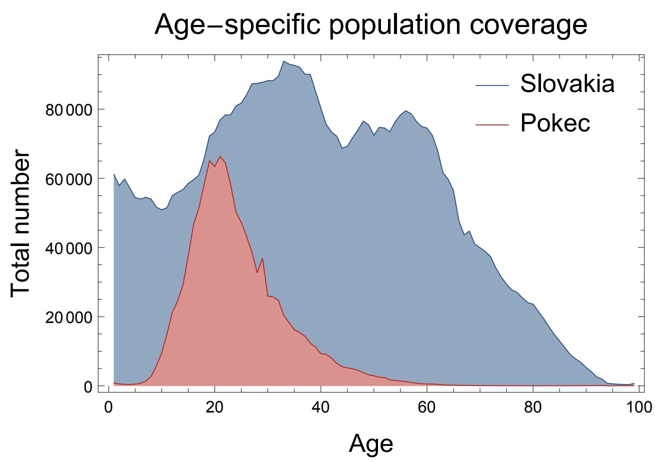

Figure 9: Coverage of Pokec users with respect to the whole population of Slovakia.

IX. DELAYED EDGE FORMATION IN OSN LAYER

To test our assumption of instantaneous link formation, we performed the following experiment. Instead of assuming the instantaneous existence of a link when both of its end nodes exist in the OSN layer, we delay its formation in the sense that we create them with a rate ξ from this point on. The rate ξ clearly has to be larger than δ, which is the rate of becoming inactive. In Fig. 10 we present the results for $\xi = 10\delta$ and $\xi = 50\delta$, which we compare with our model ($\xi \to \infty$). We observe that the position of the critical point is barely affected by the edge creation delay (see Fig. 10A and C). The initial increase in the clustering coefficient is shifted to the slightly larger network times, however, for $\eta = 0$ it reaches values similar to the case of instantaneous link formation (see Fig. 10B). The same tendency is observed for $\eta = -0.65$ (see Fig. 10C). It is important to note that the assumption of instantaneous link creation is used for the Pokec network and the model consistently. To sum up, if we assume that links are created at a timescale which is significantly smaller that the timescale at which users stop to use the network, our approximation works fine.

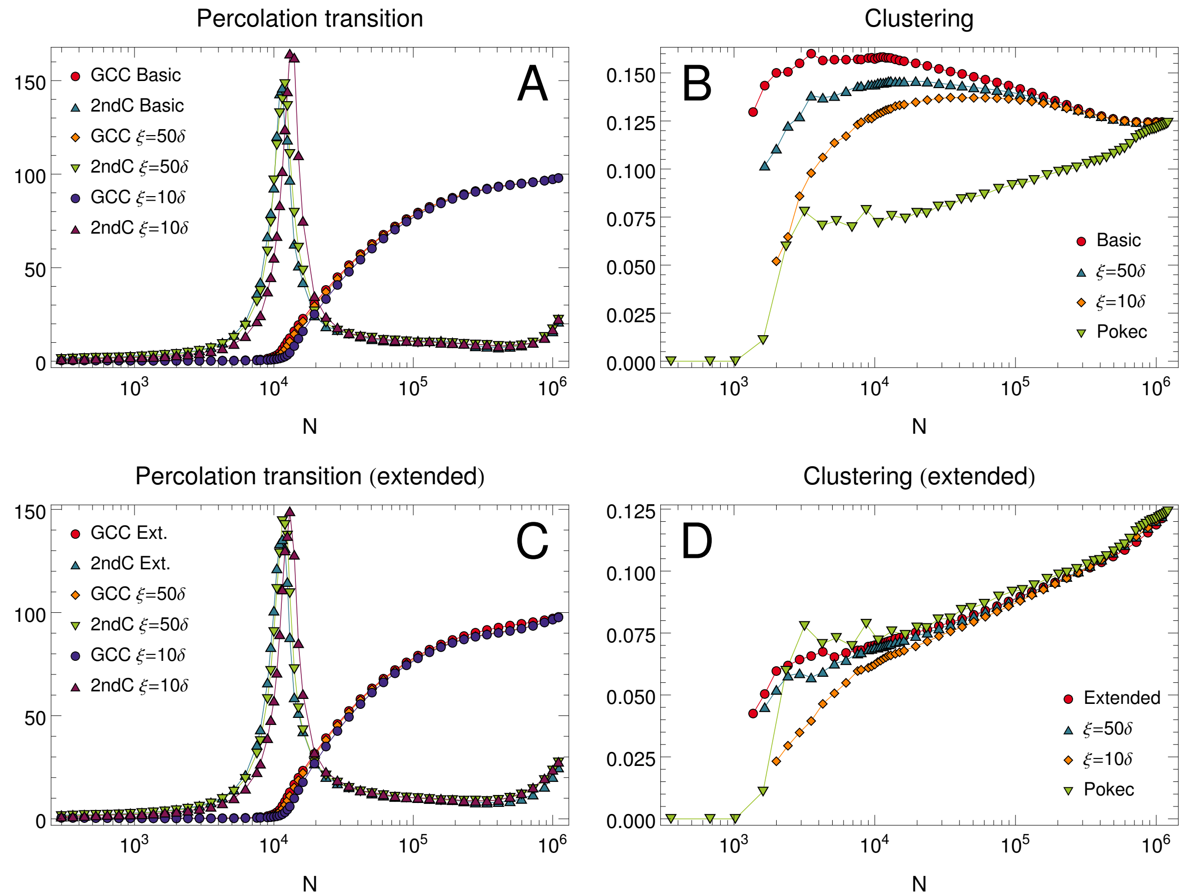

Figure 10: Results for delayed edge creation for rate $\xi = 10\delta$ and $\xi = 50\delta$. **A:** Relative size of the GCC and absolute size of the second largest component for the basic model. **B:** Mean local clustering coefficient for the basic model. **C:** Mean local clustering coefficient for the extended model. **D:** Number of components of size $N > 1$.