Supporting Appendix

Preparing the Data

The data used in this study was obtained from a mobile phone operator, from now on referred to as the "operator". We focus exclusively on voice calls, filtering out all other services, such as voice mail, data calls, text messages, chat, and operator calls. For the purpose of retaining customer anonymity, each subscription is identified by a surrogate key such that it is not possible to recover the actual phone numbers from it. Since there is no other information available for identifying or locating customers, this guarantees that their privacy is respected. We have filtered out calls that involve other operators, incoming or outgoing, keeping only those transactions in which the calling and receiving subscription is governed by the operator. This filtering is needed to eliminate the bias between the operator and other mobile service providers as we have a full access to the customers of the operator, but only partial access to the activity of other providers.

A small fraction of the subscriptions appears to be used for business or business-like purposes, which appear as users with a very large number of calls never returned. To ensure that we are dealing with genuine social interactions, we require links to represent reciprocal calls within the investigated time period, so that *A* needs to call *B* and vice versa for a link to be placed between them. This restriction eliminates telemarketing-type calls and wrong numbers. It is possible that this induces some false negatives, i.e. some links corresponding to genuine social interaction may go undetected. However, since the monitored time window is relatively long, over one third of a year, there is plenty of time to reciprocate the calls, limiting the number of false negatives.

Two quantities could be used as tie strengths: the total number and the total duration of calls placed within the period. As expected, these two variables are statistically dependent, giving rise to Pearson's linear correlation coefficient of 0.70. We have chosen to use call durations as weights (or tie strengths) w_{ij}, since they implicate the temporal and financial commitment (billing is based on call duration) to the relationship. In addition, since call durations are measured in seconds, they can be considered a continuous weight variable, whereas the number of calls suffers from strong discretization.

Given the way we have constructed the network, an interaction, or link, corresponds to a *social association* between two individuals and it is by nature bi-directional. It would be possible to retain directions in the network using directed links and thus have asymmetric weights, i.e. $w_{ij} \neq w_{ji}$, which would carry information about the distribution of calls between any two connected individuals. Yet, given that there is no *a priori* reason to assume that the individual responsible for initiating the call should interact more strongly (after all, both can interact for exactly the same call duration), we have neglected the directed nature of the links.

We allowed for the possibility that there are some very short calls which, when mapped to links, could affect the overall topology of the network. To see this, we filtered out links with total call duration less than 10 seconds per link over the examined period of 18 weeks. After this we filtered out nodes with strengths less than 60 seconds per node over the period, such that if a node is filtered out, the links connected to it are also removed. These extremely short calls do not in general represent true phone numbers, but rather mobile phone and service updates. Indeed, a common way of obtaining a new handset is by signing up for a new service, and after its activation the users switch back to the old number. We call a network without the reciprocity requirement a *non-mutual network*, i.e. a one-directional call between *A* and *B* is sufficient for them to be linked together. In contrast, a network in which the calls are required to be reciprocal is called a *mutual network*. The results of these filterings are shown in Table 1. It turns out that imposing the reciprocity condition does eliminate some of the outliers, which can be best seen in the degree distribution plots (Fig. 5). However, filtering seems to have little effect on any of the studied distributions. Consequently, in this study we used a mutual network constructed from unfiltered data.

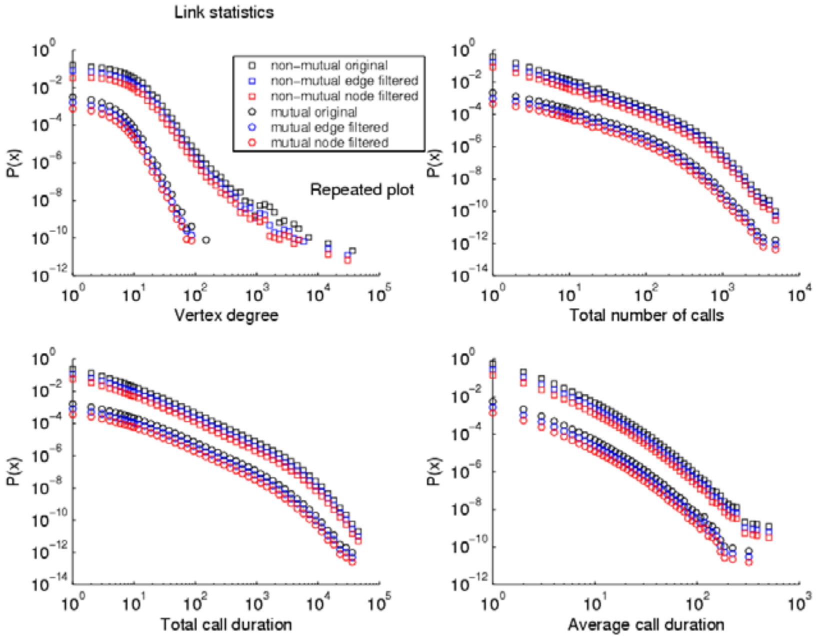

Fig. 6. Link weight distributions for mutual and non-mutual networks under different filterings. Note that the top left plot is the same as in the previous figure.

Weak ties conjecture

A direct conjecture of the weak ties hypothesis is that communities are locally connected by single weak ties (1). Granovetter justifies this conjecture by framing the hypothesis more precisely in order to derive its implications for larger networks: “The triad which is most unlikely to occur, under the hypothesis stated above, is that in which *A* and *B* are strongly linked, *A* has a strong tie to some friend *C*, but the tie between *C* and *B* is absent.” Assuming that this structure never happens, he arrives at the conjecture.

We can obtain the conjecture also using slightly different reasoning. Assume that there is a single tie *A*-*B*, known as a local bridge, connecting two communities and assume that it is strong. Based on the weak ties hypothesis, we expect the neighborhoods of *A* and *B* (which we assume exist) to overlap. But this means that there is another local bridge, a path of length 2, connecting the nodes and, thus, the two communities. This contradicts our assumption about there being just one strong local bridge and, therefore, the bridge must be a weak tie.

Sampling

The mobile phone call records, from which the network is constructed, were obtained from a major mobile operator with a market share of approximately 20% in the target country. Although the dataset covers some seven million users, it is nervertheless a sample of the underlying phone call network that consists of all mobile users in the country. In this section we investigate the possible bias of having a finite sample of the underlying network on the result shown in Fig. 1D, i.e., the increase of overlap $\langle O \rangle_w$ as a function of (cumulative) tie strength. We use the term *sample MCG* to denote the network studied in this paper and *population MCG* to denote the entire mobile phone communication network.

Let $p = 0.20$ denote the 20% market share of the operator in the target country. We assume that the nodes are all identical and that the probability of a node being governed by the operator is independent of the probability of its neighbour being governed by the operator. Given these assumptions, we can interpret p as the probability of a randomly chosen node being governed by the operator and, consequently, its being included in the sample. If we use N to denote the number of nodes in the sample MCG, the expected number of nodes in the population MCG is given by $\hat{N} = N / p = 5N$. Similarly, given the above assumptions, the probability for a link in the population MCG to be included in the sample MCG is p^2, whereas the probability for a triangle in the population MCG to be included in the sample MCG is p^3. Based on the observed sample, the expected number of links and triangles in the population MCG are, therefore, $\hat{L} = L / p^2 = 25L$ and $\hat{T} = T / p^3 = 125T$, respectively, meaning that we would expect the population MCG to contain 25 times the number of links and 125 times the number of triangles present in the sample MCG.

Since the value of p affects the number of observed nodes, links, and triangles in the sample, it is important to consider how it may affect overlap, defined in the text as $O_{ij} = n_{ij} / \lfloor k_i + k_j - n_{ij} - 2 \rfloor$, where n_{ij} is the number of common neighbors of v_i and v_j, i.e., the number of triangles around the link (v_i,v_j), and k_i (k_j) denotes the degree of node v_i (v_j). Of particular importance is the behavior of overlap averaged over links of a given weight as shown in Fig. 1D, denoted with $\langle O | w^D \rangle \equiv \langle O \rangle_w$, where the superscript in w^D emphasizes that we are using aggregated call durations as link weights. To estimate the effect of p on $\langle O | w^D \rangle$, we generate a *resample* by including each node in the LCC (largest connected component) of the sample MCG with a probability *p*. In this sampling scheme, varying probability *p* results in different sample sizes, and in the limit of setting $p = 1$ we recover the sample MCG. We consider only the LCC of the resulting resample, since for $p < 1$ the network is likely to become fragmented. The motivation for using this sampling procedure is that it mimics the way in which the sample MCG is obtained from the (unobserved) population MCG.

We chose to use $p = 0.8$, $p = 0.6$, and $p = 0.4$, and extracted three samples for each, resulting in a total of nine different samples with average sample sizes of

$\langle N_{LCC,\,p=0.8}\rangle \approx 2.6\times10^{6}$, $\langle N_{LCC,\,p=0.6}\rangle \approx 1.4\times10^{6}$, and $\langle N_{LCC,\,p=0.4}\rangle \approx 0.4\times10^{6}$ corresponding, respectively, to the different values of p. Of these the $p=0.4$ case is most interesting: the LCC of the sample MCG contains about 4.0×10^{6} nodes, roughly 10% of the estimated 35×10^{6} mobile phone users in the country, while using using $p=0.4$ results in a resample of $\langle N_{LCC,\,p=0.4}\rangle \approx 0.4\times10^{6}$ nodes, roughly 10% of the nodes in the LCC of the sample MCG. The results are shown in Fig. 7. Although lower values of p result in slightly lower values of $\langle O \vert w^{D}\rangle$, its qualitative behavior is fairly insensitive to p, and the curves have the same characteristic features as the one in Fig. 1D. Further, examining average overlap as a function of cumulative weight (Fig. 7B) shows that the curves become slightly steeper as p increases. Consequently, it is safe to assume that the behaviour of $\langle O\rangle_{w}$ is unaffected by the finite sample. Had we access to the records of all mobile phone users in the country and not just those of a single operator, based on Fig. 7B, we would expect an even more pronounced increasing trend for $\langle O\rangle_{w}$.

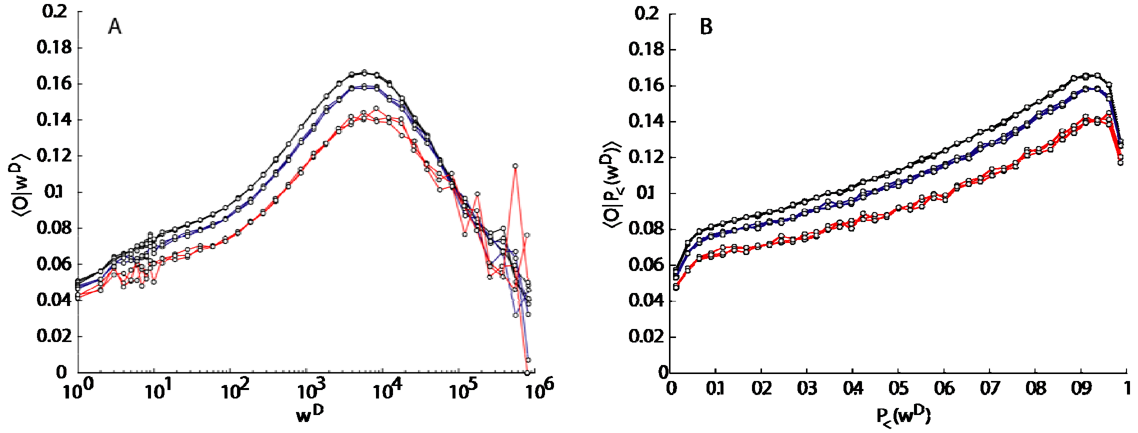

Fig. 7. (A) Average link overlap $\langle O \vert w^{D}\rangle$ as a function of absolute weight, the aggregated call duration w^{D}, and **(B)** average link overlap $\langle O \vert P_{<}(w^{D})\rangle$ as a function of cumulative weight $P_{<}(w^{D})$, corresponding to the fraction of links with weight less than or equal to w^{D}, for different network samples. There are altogether nine curves in each plot, corresponding to three different samples for each of the three chosen values of extraction probability $p=0.8$ (black), $p=0.6$ (blue), and $p=0.4$ (red). The curves corresponding to different samples for a fixed value of p practically coincide. While lower values of p result in slightly lower values for the average overlap, the qualitative behavior of the curves remains unchanged. This demonstrates that the result of Fig. 1D, i.e., the higher the value of w^{D} the higher the value of overlap O on average, is not sensitive to having a one-operator-sample of the underlying phone network, and it can be reproduced for sub-samples of the original sample (original data).

Interdependence of Weights and Topology

Fig. 8A shows overlap O_{ij} averaged over all links with weight w, as a function of weight w, indicating that while for small weights ($w < 10^4$) the overlap $\langle O \rangle_w$ increases with w as expected, for large weights ($w > 10^4$) the overlap $\langle O \rangle_w$ actually decreases. This means that in the region above $w \approx 10^4$, the stronger the tie, the smaller the overlap. This surprising decreasing trend can be understood by considering the link weight distribution shown in Fig. 8B, from which we find that only about 5% of the links lie in the w>10^4 region. To correct for this uneven weight distribution in the paper, we plot the overlap as a function of the cumulative link weight $P_{\text{cum}}(w)$, which is the percentage of links with weight smaller than w, and it is shown in Fig. 1D in the paper.

Since the decreasing trend for top 5% of weights concerns some 325 000 links, it cannot possibly be attributed to insufficient statistics. The links in this region correspond to pairs of users who devote more than three hours to each other over the investigated period. Our measurements indicate, however, that they have a common property: These individuals devote the vast majority of their on-air time to a single acquaintance, and the time spent with others is negligible. Consider a link located between vertices v_i and v_j carrying weight w_{ij}, and denote the strengths of the adjacent nodes with s_i and s_j, respectively, defined as $s_i = \sum_{j, j \in N(v_i)} w_{ij}$, where the sum index j runs over the neighbours of node i. The smaller of the strengths is given by $\min(s_i, s_j)$ and the larger by $\max(s_i, s_j)$, unless the strengths are equal. The ratios $\min(s_i, s_j)/w_{ij}$ and $\max(s_i, s_j)/w_{ij}$, shown in Fig. 8C, correspond to the strengths of the nodes measured in units of the link weight w_{ij}. For weak links (small w_{ij}) both of these values are high, meaning that overall both adjacent nodes spend a considerably longer time on the phone than they do talking to each other and, thus, the link connecting them constitutes only a small fraction of their on-air time. As we move towards strong links (high w_{ij}), we find both ratios decreasing and eventually converging to one at approximately w=10^4. This demonstrates that for strong links, in the region where $\langle O \rangle_w$ start to decrease in Fig. 8A, the strengths of both adjacent nodes are about as large as the link weight w_{ij} and, thus, the high weight relationship clearly dominates the on-air time of both users. Consequently, both have less time to interact with other acquaintances, explaining the onset of the decreasing trend for $\langle O \rangle_w$ in Fig.8A.

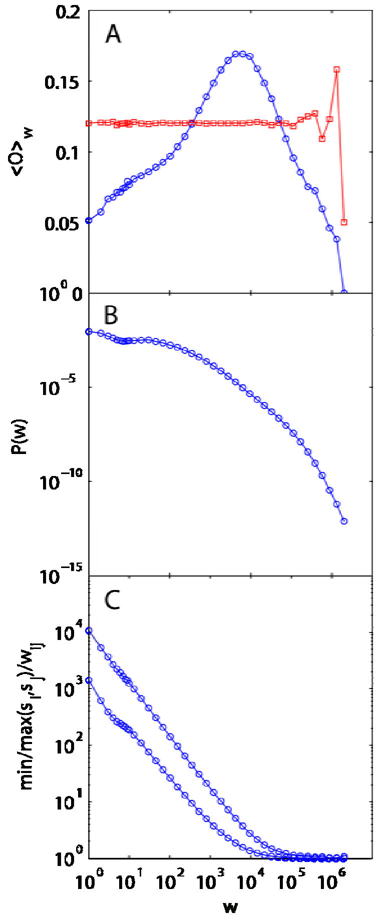

Fig. 8. (A) The overlap of link neighborhood $\langle O \rangle_w$ increases as a function of the link weight *w* (blue circles) up-to $w \approx 10^4$, revealing a statistical connection between local network topology and link weights. A random reference (red squares) is obtained by randomly permuting the weights, thus removing the coupling between $\langle O \rangle$ and *w*. Surprisingly, for large weights $w \approx 10^4$ the overlap $\langle O \rangle_w$ actually decreases in this region, apparently contradicting the weak ties hypothesis. Yet, as we explain, that region represents a minority of the users. **(B)** The distribution of links weights w_{ij} decays fast, with only 4.4% mass to the right of $w = 10^4$. This means that the decreasing part of the O_{ij} curve applies to less than 5% of links, which is seen clearly by plotting the overlap O_{ij} as a function of cumulative weight $P_{\text{cum}}(w)$ as in Fig. 1D. **(C)** The fraction of total time (node strength) devoted by the adjacent nodes to a link of weight w_{ij} is given by $\min(s_i, s_j)/w_{ij}$ and $\max(s_i, s_j)/w_{ij}$, and is here plotted as a function of weight *w*. Values close to one indicate that the communication is almost entirely focused on one individual in the $w \approx 10^4$ region.

Betweenness Centrality for Links

For a link $e=(v_i,v_j)$ we can write betweenness centrality b_{ij} as

$$b_{ij} \equiv \sum_{v\in V_s} \sum_{w\in V/\{v\}} \frac{\sigma_{vw}(e)}{\sigma_{vw}} \qquad (2)$$

where $\sigma_{vw}(e)$ is the number of shortest paths between v_v and v_w that contain e, and σ_{vw} is the total number of shortest paths between v_v and v_w (2). In practice, we use the algorithm introduced in (3) to compute b_{ij} but, due to limited computing capacity, instead of using all the nodes of the set V making up the network, we use a subset 10^5 nodes in the sample $v\in V_s$ as *starting points*. The size of the set V_s is given by N_s.

Determining the nature and the position of the phase transition point

The transitions observed in Fig. 3 suggest two important questions: How does the position of the critical threshold p_c depend on the size of the system? Are the transitions genuine phase transitions or finite size effects? In order to answer these questions, we carried out finite size scaling (FSS) for all four different thresholding schemes (remove min w_{ij}, min O_{ij}, max w_{ij} and max O_{ij} links).

In many large systems studied by statistical mechanics, from gases to magnetic materials, the system is considered infinite in the number of its constituent elements. Different quantities of interest can be expressed in terms of the correlation (connectivity) length ξ of the system, which in the vicinity of a phase transition diverges like $\xi\sim|p-p_c|^{-\nu}$, where ν is the critical exponent for correlation length. However, in a finite system, the correlation length is limited by the system size, and the divergence becomes rounded. Consequently, other quantities related to the correlation length also show a rounded signature of the divergence, but never actually diverge due to finite N (4), as demonstrated for path length in Fig. 9. In general, the location of the transition p_c depends on the system size as

$$|\, p_c(N) - p_c(\infty)\,| \sim N^{-\chi}, \qquad (3)$$

where $p_c(\infty)$ corresponds to the extrapolated value in the thermodynamic limit as $N\to\infty$. Here the value of the exponent χ is related to ν and it quantifies how changing the system size affects the position of the critical threshold, whereas the extrapolated value $p_c(\infty)$ reveals the nature of the transition: If $0 < p_c(\infty) < 1$ there is a real phase transition but, on the other hand, if $p_c(\infty) = 1$ there is no actual phase transition, and the observed signature is caused by the finiteness of the system (5).

We use two different sampling techniques to produce systems of different size *N*. In the first approach, we choose randomly one node in the initial network as a source node v_s and extract a radius ℓ neighborhood of this node, including all nodes, and links between them, which are at most a distance ℓ from v_s (less than ℓ links from v_s). The size of the sample depends exponentially on the extraction depth ℓ, so that increasing the value of ℓ enables us to extract larger samples, although the realized sample size *N* will depend on the starting node v_s due to the non-homogeneous topology of the network. We call such a sample an *extract*. In the second approach, we generate a sample of the original network by mimicking the process responsible for generating the network. Consider the original network to represent the true underlying social network of which we see only a part, since phone calls are just one form of social interaction. We then assign an occupation probability *p* to each node in the network, corresponding to the probability that this node is governed by the operator and, thus, belongs to our sample. This means that the probability for a given node to belong to the sample, i.e. to be occupied, is an independent random trial and does not depend on whether its neighbors are occupied. In this sampling scheme varying the node occupation probability *p* results in different sample sizes, just as varying extraction depth ℓ does in the extract sampling scheme. The original network corresponds to *p*=1, whereas if *p*<1 the network is likely to become fragmented, in which case we take the largest connected component (LCC) as our sample. A sample obtained using this second method is called a *resample*.

To carry out FSS we need to know the size of the system *N* and the location of the transition $p_c(N)$ for this finite system. Having several $N, p_c(N)$ point pairs, corresponding to different system sizes, allows us to extrapolate the value of $p_c(\infty)$. The system size *N* is just the number of nodes in the given sample and can be obtained easily, but finding the value of $p_c(N)$ is a bit more laborious in practice. In principle, we can find this from the behavior of susceptibility, defined as $S = \sum_s n_s s^2 / \sum_s n_s s$, where n_s is the number of clusters, per lattice site, containing *s* sites and the LCC is excluded from the sum, but in practice we measure $\sum_s n_s s^2$, which behaves similarly to *S*. Although *S* diverges for *N*→∞ at the transition point, in practice it is rather noisy even for medium size systems, making it difficult to pinpoint the location precisely. A more robust technique is to use the smoother, monotonically decreasing order parameter, defined as the fraction of nodes in the LCC and written as $R_{LCC} = \sum_s n_s s$. R_{LCC} is expected to vary most rapidly at the threshold - in fact $\partial R_{LCC} / \partial f$ usually diverges in an infinite system. We can find the location of the transition by computing this derivative numerically and by identifying its steepest descend point with the transition point, which should coincide with the point of divergence for susceptibility. This method works better, but the numerical derivative is not sufficiently robust for smaller systems.

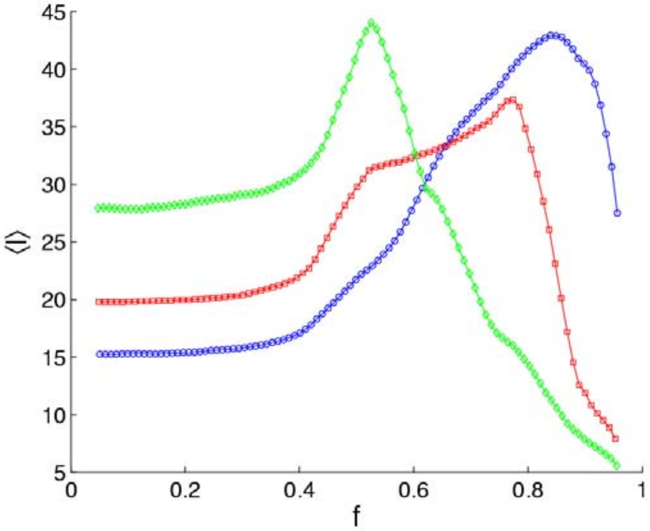

Fig. 9. Rounded signature of divergence of average shortest path length $\langle \ell \rangle$ due to finite system size. The green (◊), red (□), and blue (o) curves are associated with system sizes $N \approx 4.4\times10^5$, $N \approx 1.4\times10^6$, and $N \approx 3.3\times10^6$, and they were obtained using resampling extraction with node occupation probabilities *p*=0.40, *p*=0.60, and *p*=0.90, respectively. For each value of node occupation probability *p* we sampled a few systems, carried out the thresholding for each of them, computed the average $\langle \ell \rangle$, and then finally smoothened the plot with a moving window average. The critical point f_c moves to the right as the size of the system increases and, since there is no phase transition in this case, we have $f_c \rightarrow 1$ as $N \rightarrow \infty$. Note that the starting value of $\langle \ell(f=0) \rangle$ is highest for the smallest system, because the networks are more tree like there as, relatively speaking, more links are missing from small than large samples, suppressing the small world effect of short paths.

Fortunately, Eq. 3 is valid for every reasonable definition of a percolation threshold for finite large systems, and it is only the proportionality constant that is different for different definitions of the onset of percolation (5). It turns out that in this case the most reliable results are obtained by manually determining the transition point $p_c(N)$, denoted in the text with $f_c(N)$, from plots of order parameter R_{LCC} vs. the control parameter *f*. We then make a of plot $f_c(N)$ vs. 1/*N* and fit, in the sense of least sum of squared error, a second order polynomial to the data. The transition point in the infinite size limit is extrapolated from the *y*-intercept of the fit. In some cases the coefficient of the second order term is close to zero, so that the fit effectively is linear. In most cases, however, the fit is clearly curved, indicating that the exponent *x* is different from -1. Theoretically, the inclusion of the second order term can be justified as a correction to the leading scaling behavior. The correction vanishes as $f \rightarrow f_c$, but its contribution may be significant even if $| f - f_c |$ is small. We use this method to obtain sets of estimates of $f_c(\infty)$ for different thresholding schemes and sampling techniques using bootstrapping, in which we randomly choose half of the points to be included in the bootstrap sample, and find out the value of $f_c(\infty)$ using only the points in the bootstrap sample. Repeating this 10000 times gives a distribution of the estimates of $f_c(\infty)$. We take the value $f_c(\infty)$ as the mean of the bootstrap distribution, and the error bounds are taken as the standard deviations of the $f_c(\infty)$ distribution (6).

The results of finite size scaling are given in Table 2. From the extrapolated values it is clear that we have phase transitions as $f_c(\infty)\neq 1$ for descending thresholding, and the values are different for weight-driven and overlap-driven thresholding schemes. The intervals of plausible $f_c(\infty)$ values, taken as the region that is one standard deviation from the mean, are slightly different for extracts and resamples, but part of these intervals coincide. Both sampling techniques support the idea that $f_c(\infty)$ for descending weight thresholding lies in [0.83,0.84], while $f_c(\infty)$ for descending overlap thresholding lies in [0.66,0.67]. Put together, these result fully support the existence of a phase transition for these thresholding schemes.

The results for ascending thresholding are not quite as clear. The coincidence intervals of extracts and resamples for weight and overlap thresholding are [0.89,0.92] and [0.93,0.97], respectively. In the latter case of ascending overlap thresholding, $f_c(\infty) = 1$ is contained within one std of the mean for resamples and within two stds of the mean for extracts, suggesting that instead of a phase transition we most likely have a finite size effect. For ascending weight-driven thresholding $f_c(\infty) = 1$ is not included within two standard deviations of the mean. However, there are two important practical aspects to be kept in mind when interpreting the results for ascending thresholding. First, the behavior of the order parameter R_{LCC} as a function of the control parameter f is noisier for ascending than descending thresholding, with the effect that estimating the finite thresholds $f_c(N)$ is more prone to errors in the ascending scheme. This is a consequence of the structural properties of the studied network. Second, the manual estimates of $f_c(N)$ may have a slight downward bias. Since $f_c(N)$ must lie in the [0,1] interval, one would not estimate $f_c(N) > 1$ for any sample as this does not have any physical meaning. Thus, we conclude that the ascending overlap-driven thresholding exhibits no phase transition; In the case of ascending weight-driven thresholding the transition point is dramatically shifted upward, and is compatible with the assumption of no transition for $f < 1$.

Overall, the network's response to removing weak links is qualitatively different from the response to removing strong links, but quite independent whether we use weights w_{ij} or overlap O_{ij}. Our results also suggest that the transition observed for removing strong links first is a finite size effect (f_c=1), whereas the transition for removing weak links first is a genuine phase transition ($f_c\neq 1$). This means that the observed qualitative difference between weak and strong links is not a consequence of using the given, finite size sample, but demonstrates that weak and strong links are qualitatively different regardless of the size of the system.

Scheme	n	$f_c(\infty)$
DW (extract)	17	0.80±0.04
DW (resample)	21	0.85±0.02
DO (extract)	17	0.62±0.05
DO (resample)	21	0.69±0.03
AW (extract)	17	0.89±0.03
AW (resample)	21	0.91±0.02
AO (extract)	17	0.92±0.05
AO (resample)	21	0.98±0.05

Table 2. A summary of finite size scaling (FSS) results. The key to the different thresholding schemes is the following: A = ascending (remove max links first), D = descending (remove min links first), W = weight driven thresholding, and O = overlap driven thresholding. The words extract and resample in parentheses refer to extracted and resampled samples, respectively, on which the FSS is based. The number of available samples, after the smallest ones were discarded, is denoted with *n*. The number of samples used in each bootstrap realization is *n*/2, and the $f_c(\infty)$ is the value of the percolation threshold extrapolated in the thermodynamic limit as $N \rightarrow \infty$.

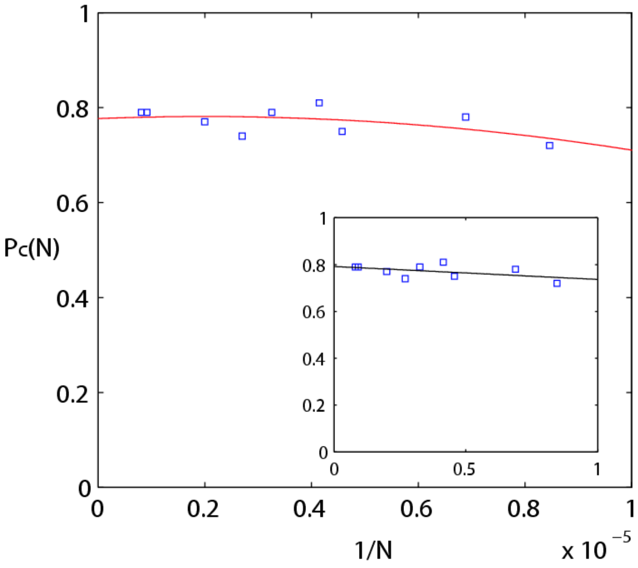

Fig. 10. Main panel: One realization of a bootstrap sample for descending weight-driven thresholding, using an extract sample, and the corresponding second order polynomial fit to it. Inset: A linear fit to the same data. Both fits yield practically identical results. The extrapolated value $p_c(\infty)$ for this

particular sample is approximately 0.78. Since $p_c(\infty)$ is clearly less than unity, this corresponds to a genuine phase transition. In the case of no transition we would have $p_c(\infty) \approx 1$.

Spreading model

Considerations of information flow lead us to formulate a simple model in which the spreading probability from an infected node v_i to its nearest susceptible (non-infected) neighbor v_j was made proportional to the link weight w_{ij}. Introducing a constant of proportionality *x*, we write the time independent probability of passing information from v_i to v_j as $P_{ij}=xw_{ij}$, where increasing *x* results in a higher spreading probability. The most obvious choice is to set $x = 1/\max_{ij}(w_{ij})$, in which case for the globally strongest link we have $P_{ij}=1$, and for all others $P_{ij}<1$. While this is a reasonable choice, it results in extremely long running times for the simulation. The reason for this is the highly skewed weight distribution *P*(*w*), so that normalizing with the globally maximum weight, which can be seen as an outlier, results in very low transmission probabilities for most links, requiring a large number of trials before any macroscopic spreading takes place. This problem is amplified by the fact that the simulations, both for empirical and random network, should be carried out for an ensemble.

We can circumvent this problem by increasing the value of *x*, which speeds up the simulations without affecting the qualitative behavior of the system and, thus, it can be seen as a rescaling of the time axis (Fig. 4, A and B). This introduces a cut-off w^* for the transmission probability P_{ij}, below which it is linear with respect to w_{ij}, and unity for $w_{ij} \geq w^*$ (Fig. 11). But how should one choose the value for *x* or, alternatively, for the cut-off w^*? While the location is to some extent arbitrary, a range of values suggests itself using the following reasoning. For choosing a suitable value for w^*, let us deal in terms of the cumulative weight distribution $P_{cum}(w)$, and choose a value for $P_{cum}(w^*)$ instead. The first requirement is that the relationship $P_{ij} \sim w_{ij}$ should be valid for at least half of the links, since otherwise we can hardly say that the two are proportional, and this gives us a lower limit $P_{cum}(w^*) > 0.5$. Since we are interested in the effect of the coupling between weights and topology on a dynamic process, we will stick to a region of link weights in which this observed coupling holds, and from Fig. 1D we see that this is the case up to $P_{cum}(w) \approx 0.95$, giving an upper limit of $P_{cum}(w^*) \leq 0.95$. Within the lower and upper limits, we would like to have as high a value of $P_{cum}(w^*)$ as possible, but also to ensure that we stay away from the region with anomalous behavior (overlap decreasing as a function of weight, a phenomenon that may be specific to the mobile phone network). These heuristics lead us to choose $P_{cum}(w^*) = 0.90$, which for the studied period of 18 weeks corresponds to $w^* = 3867$ seconds, i.e. a little over an hour, or $x = 1/w^* \approx 2.59 \times 10^{-4}$ 1/s. With this choice, the intended relationship $P_{ij} \sim w_{ij}$ holds for 90% of the links, while for the strongest 10% of the links the transmission always takes place. It turns out, however, that the qualitative nature of the spreading results is fairly insensitive to the precise value of *x*, i.e. the weight permuted network performs better at spreading than the empirical network for different values of *x*.

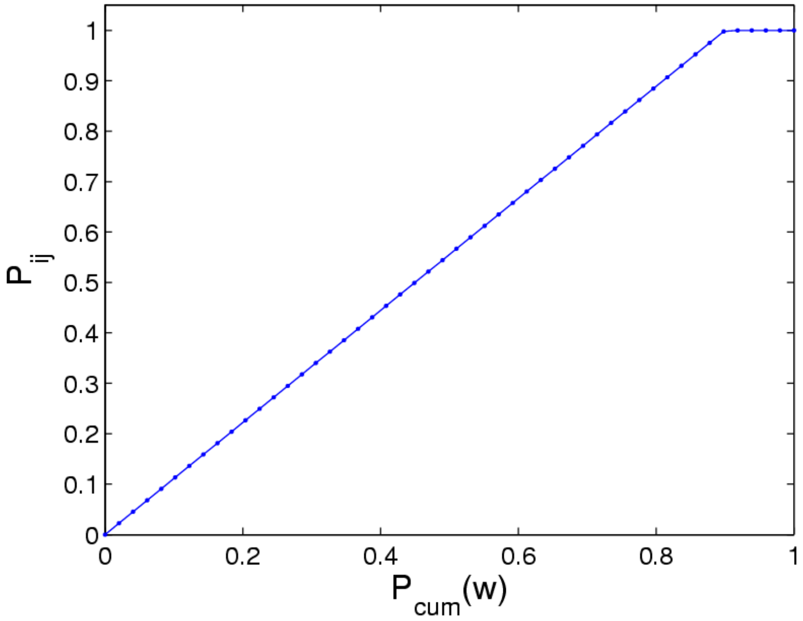

Fig. 11. The transfer probability P_{ij} as a function of cumulative link weight $P_{\text{cum}}(w)$, the fraction of links with weight less than *w*. Using the value of $x \approx 3.0 \times 10^{-4}$ results in a cut-off at $w^* \approx 0.90$, and thus the intended relationship P_{ij}~w_{ij} applies for 90% of links.

Note that although the cut-off was introduced for computational purposes, its existence may, in fact, be a desirable property. Common sense tells us that some pieces of information are more important than others or, in the context of gossip, some pieces of gossip are juicier than others. Lowering the cut-off point w^* means that we have more links with P_{ij}=1, such that if v_i has access to information, it will always pass it on to its neighbor v_j as long as $w_{ij} \geq w^*$. For lower cut-off points this will be true for an increasing number of links in the network, and soon rumors will spread like wildfire.

References and Notes

1. Granovetter M (1973) *Am J Sociol* 78: 1360-1380.

2. Holme P (2002) *Phys Rev E* 66: 36119.

3. Newman MEJ (2001) *Phys Rev E* 64: 16132.

4. Newman MEJ & Barkema GT (1999) *Monte Carlo Methods in Statistical Physics* (Oxford University Press).

5. Stauffer D & Aharony A (1994) *Introduction to Percolation Theory* (CRC Press, ed. 2).

6. Efron B & Tibshirani RJ (1994) *An Introduction to the Bootstrap* (Chapman & Hall/CRC).