Observing Communities and Creating Togetherness

by Dmitry Paranyushkin, Berlin, November 2010

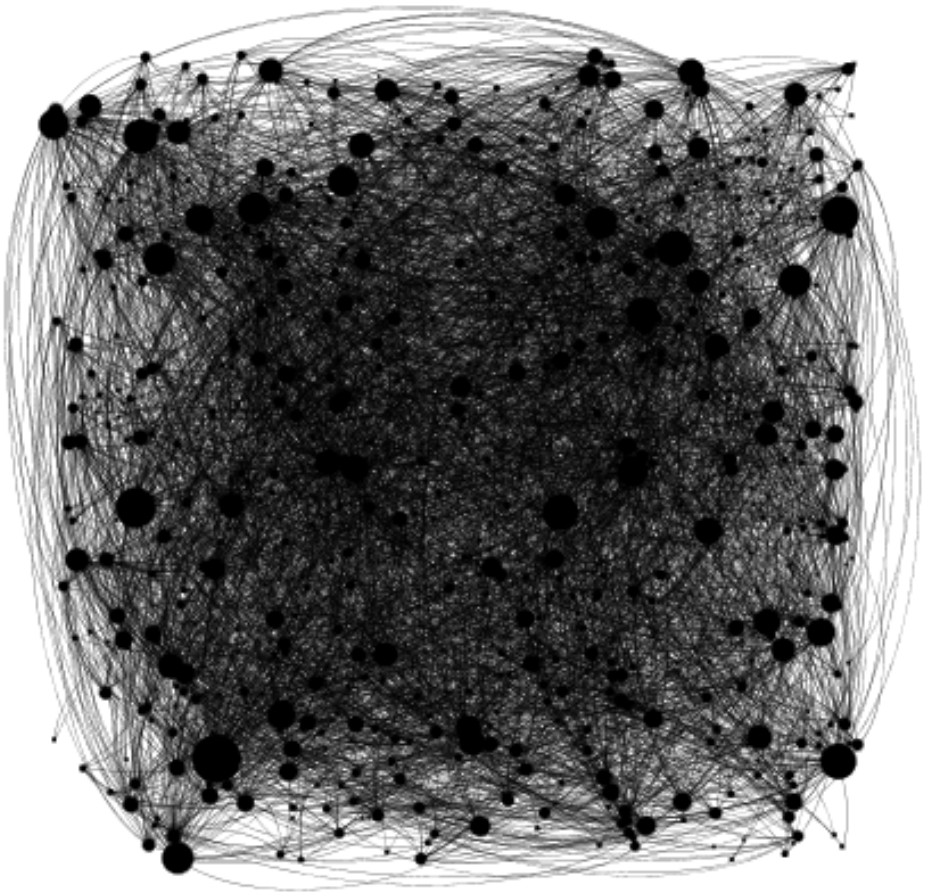
While there are many ways to think about a community, let's model it as a network. The nodes represent the elements of the community and the edges represent their interactions. The methods used in network analysis reveal the factors that bring communities together, enable them to function, and cause them to fall apart.

"A qualitative multiplicity is not an aggregate of parts with an apparent unity constituted by the relation of separate numerical or physical existents (the Galilean world of purely external relations) but an event, an actual occasion of experience". [1]

So community starts from "an occasion of experience", from an act of observing that is temporal and experiential.

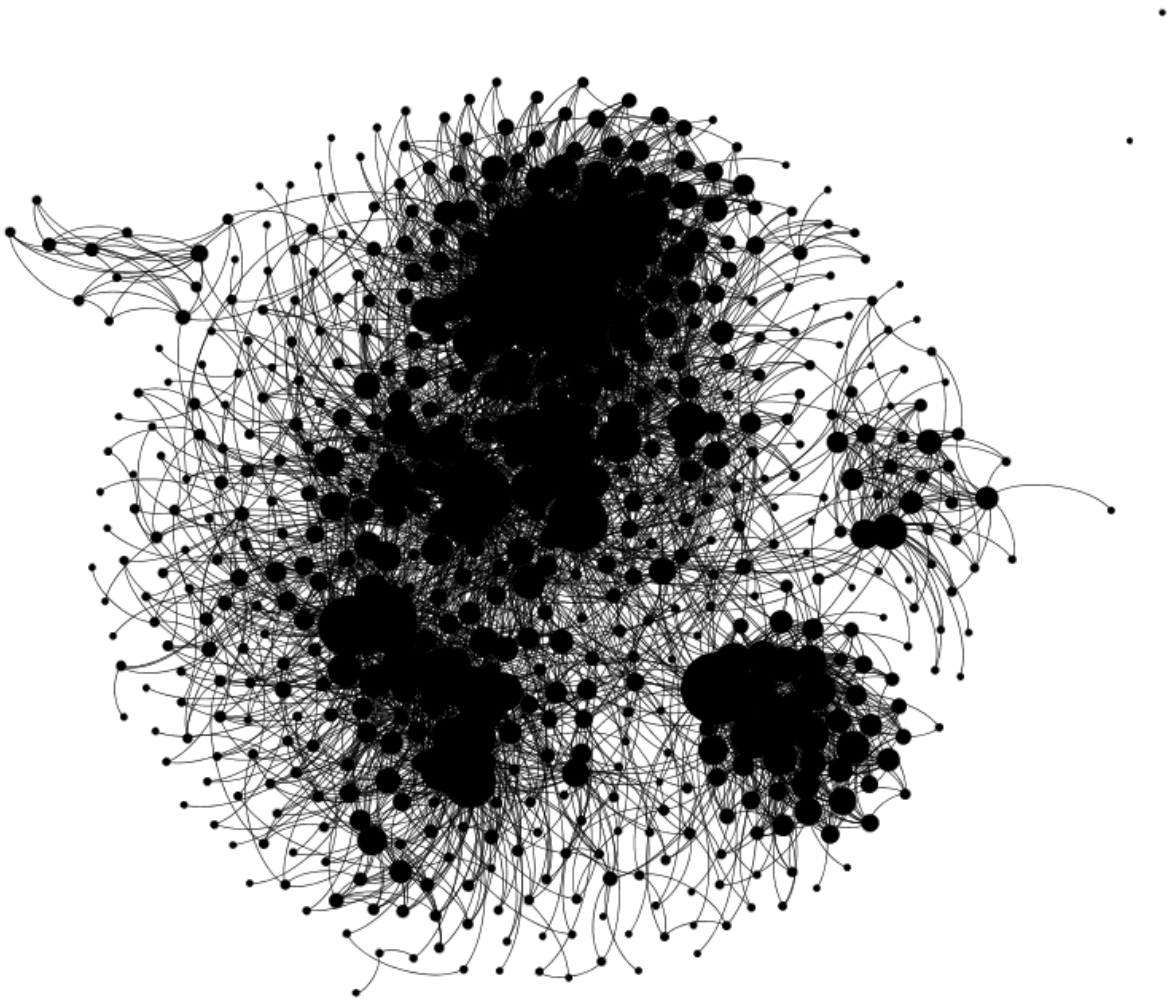
An Occasion of Experience

How would we "observe" a community modeled as a network? Each member is a node, each relation is a thread. A snapshot of their interactions would look something like the graph below: 649 Facebook 'friends' where the 3926 edges indicate their connections. The nodes that have more connections are bigger on the graph.



Here each node has a random position and every connection is assumed to have the same quality and intensity. This community looks like a swarming mass – almost no distinctions and no differences.
As mentioned before, community has to do with the act of observing. Let's borrow several graph layout techniques in order to see how our understanding of a community changes depending on how we observe it.

Some of the most interesting things we can observe about a network are the clusters or sub-communities it's comprised of. There are many algorithms to produce readable graph visualizations. Most existing "energy models" group densely connected nodes and separate sparsely connected ones, enforcing uniform edge lengths and therefore preventing the separation of nodes into different clusters [2]. As a result, nodes that have more connections tend to be positioned in the middle of the graph (Frutcherman-Reingold method):



While this visualization makes the community more "readable", it's not yet very easily "interpretable". The reason is that most of real-world networks are scale-free: the nodes form clusters around a few, but significant number of hubs. [3] The density of connections between the nodes should be taken into account, in order to be able to observe the formation of the community in a way that reflects its natural properties. An algorithm used in Gephi [4] (similar to LinLog edge repulsion algorithm proposed in [2]) brings closer the nodes that have less

connections and pushes further apart the nodes that have more connections. This way the nodes with less connections are aligned around the hubs that are dispersed over the Euclidian space of the graph.

The same community now clearly has at least four different visible sub-communities or clusters. What brings the nodes together here is the intensity of their relations. The less nodes they interact with, the closer they are to the ones they know. The hubs are pushed apart from each other, but are surrounded by the nodes they are connected with – closer to those that have fewer connections to the rest of the network. There are also some elements on the periphery, a free-floating "foam" who do not belong to the main component.

The role of the hubs is quite important, because they are the ones that spread information through the clusters they belong to. They also play an important role in connecting the network together and providing the most efficient routes for the information to proliferate through the network.

Hubs, Robustness, and Power

As can be seen on the last image, there are about 20 major hubs or people who accumulate the most connections (and power) within the network. This property of real-life social networks is called scale-free and has quite important implications on how this network functions. Scale-free networks are more robust against random node removal, however are susceptible to targeted attack on the most connected hubs [3].

What happens when the major hubs and the nodes with the highest betweenness centrality are removed from the network? Research shows that even when a seemingly insignificant number of the hubs is removed from scale-free network (~ 7% for the world wide web) it fall will apart into disconnected components [6].

If 10% of the most connected nodes are removed, the average path in the network increases by 30% of the original 3.2 to 4.3 and clustering decreases from 0.33 to 0.22. Higher average distance means that it becomes more expensive and slower to distribute information within the network. Lower clustering means that there are less informational loops in the network and it makes it more probably that the information (or disease) will spread to every member of the network [7].

The number of connections falls to 1590, it decreases almost twice: 10% of nodes were responsible for 50% of connections, thus the efficiency of communication in this network suffers even more. Also, what happened is that one community broke off from the main component – a sign of the starting fragmentation.

So, in the social network above the hubs ensure that information spreads easier and faster, however, their presence makes it also harder for a message to reach the periphery of every community, as it sometimes remains circulating among the most connected nodes of a cluster.

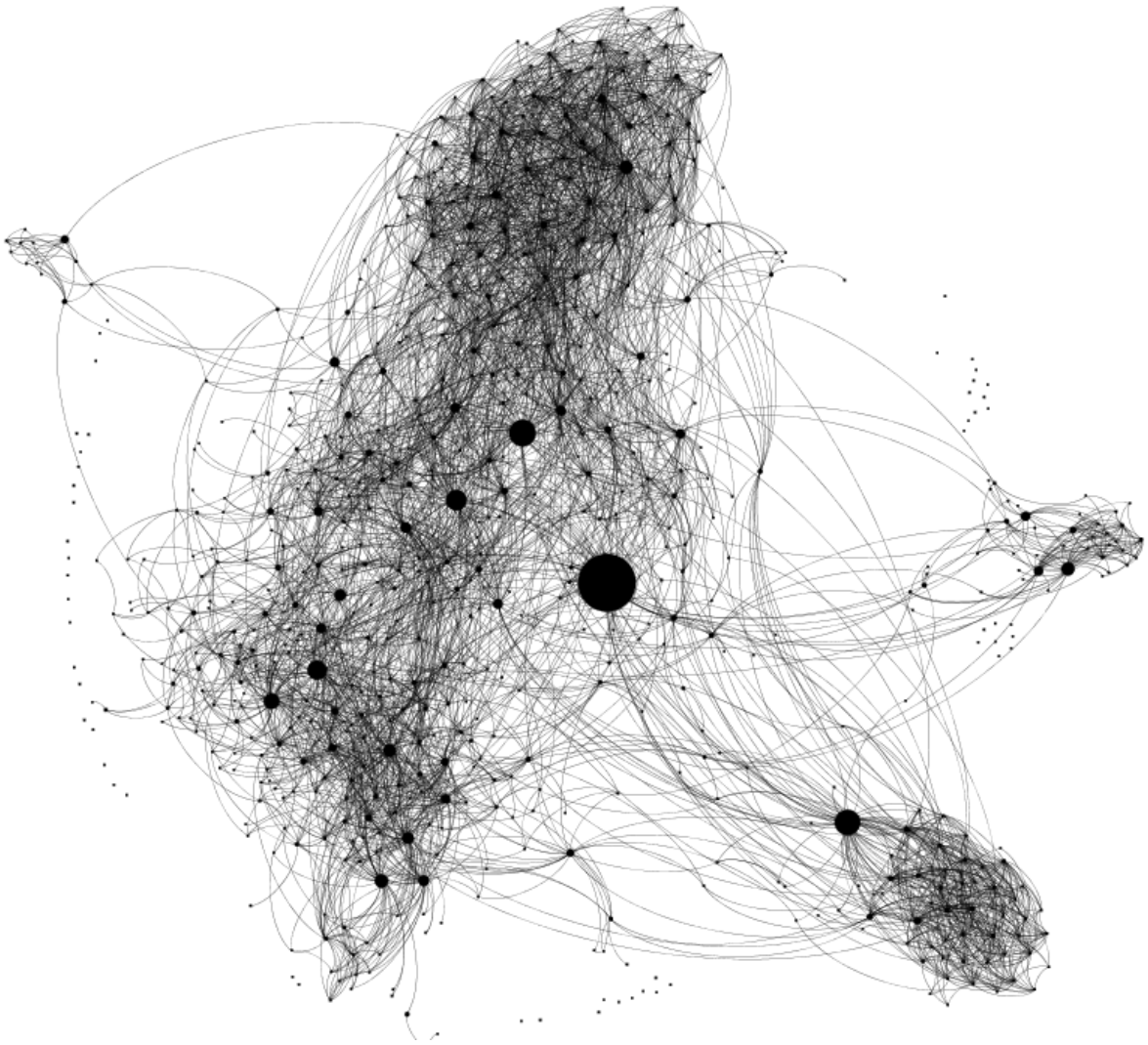
Power, or the number of connections each node has, is still distributed unequally after removing the major hubs. The degree power law in the previous network was 5.37 and in the new one it's 5.5. So inequality is not resolved by removing the most connected hubs.
The reason is that in such a large network as soon as the hubs are removed there are the new ones taking on their functions. The network has designated roles and everyone is replaceable, so it's not the hubs but the structural principle of the networks (smaller nodes located around the most connected hubs) that proliferates the power principle inside.

Overall, the hubs play an important rule in bringing the community together. There are only a few of them, but they are in control of the most informational flows within the network. They ensure that information is distributed within the network efficiently, through the shortest pathways. At the same time, their presence makes it more difficult to ensure that a certain message will reach every member of the community (especially the periphery), because hubs create clusters (or sub-communities) within the community, which makes it harder for information to break out to the whole network at a large. This goes in line with the findings that random networks synchronize more readily than scale-free ones [8]. In random networks there are no hubs, each member has more or less the same number of connections and influence, so depending on the level of interconnectedness within the network a certain message might mobilize the whole network.

If a scale-free network introduces more randomness inside its organizing principle, it will be easier to mobilize. At the same time, this process should be done carefully, because at some point, as the number of connections grows, the network may become unstable and too susceptible to disease propagation and uncontrolled oscillation. It is therefore useful to introduce a certain time limit on the life-span of these random connections in order to keep their number in check. [9]

Centrality and Shortcuts

There are also other ways to look at this community. For example, a very interesting measure is Betweenness Centrality, which shows how often a node appears on the shortest path between any two random nodes in the network, and therefore its potential for control of communication [5]. People who have the highest betweenness centrality act as the linchpins within the community, forming its backbone. They are usually the ones who hold the various groups together and make the community what it is. They are also the ones who can spread the message not only within their group, but within the whole community at large. They are the shortcuts for information that belong to a certain cluster and are responsible for connecting that cluster in the most efficient way to the rest of the network.



It can be seen that there are about 6 people in the network with their betweenness centrality significantly higher than the rest of the network within this community.

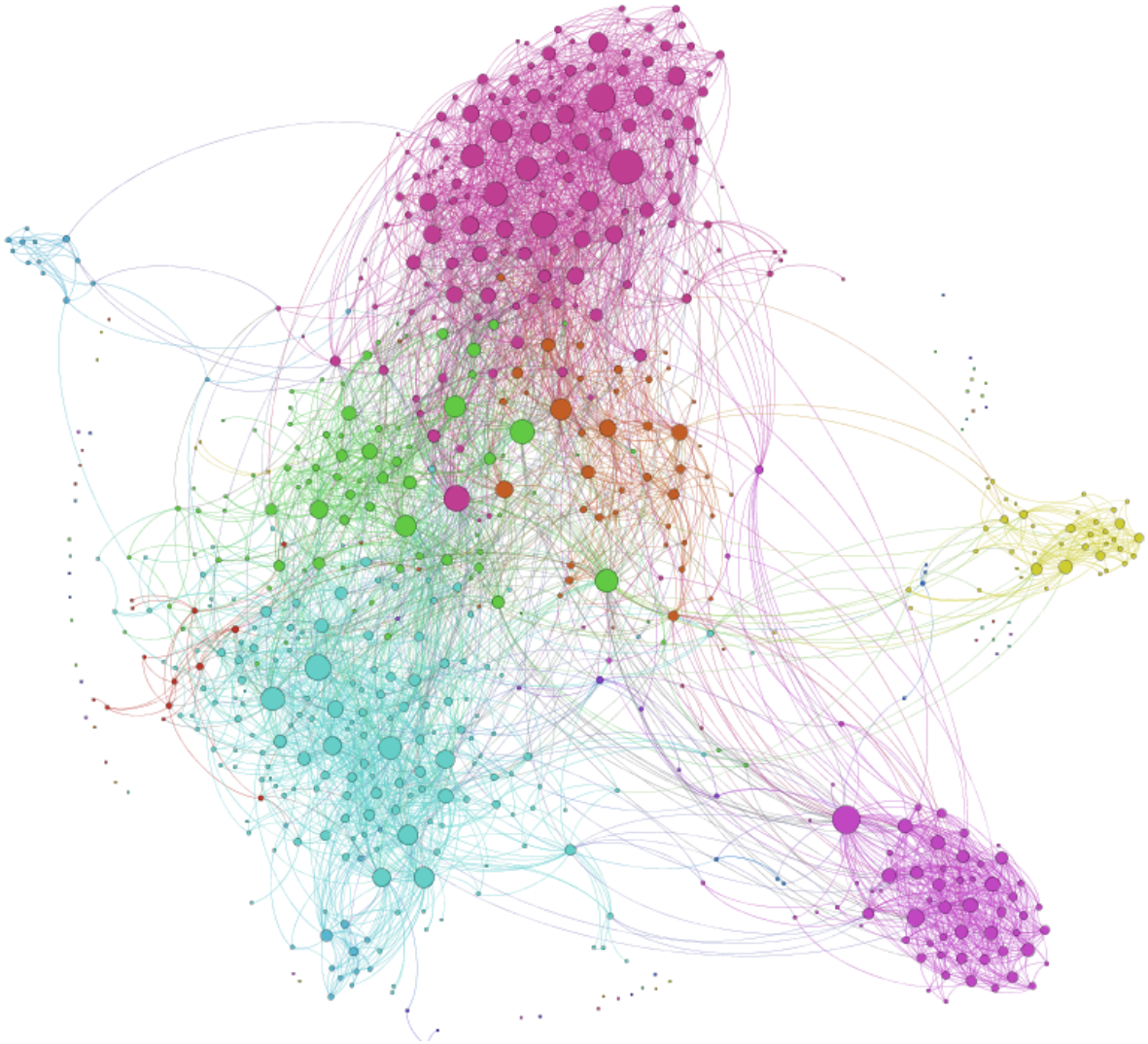
What's interesting is that when we start removing the nodes with the highest betweenness centrality, t's not the next at the queue who are getting the most influence for this social network. Sometimes removing an important connecting linchpin hub can increase the importance of a node that wasn't as significant before. This result can prove to be very interesting for analyzing political and criminal networks in order to see what new centers arise with the removal of central nodes.

To summarize, an efficient strategy in bringing together a community is transforming it from scale free to random – introducing more random links between the less connected nodes, within the periphery. This should be done in order to retain the efficiency of communication that the hubs provide, while also making it possible for the network to synchronize and, thus, mobilize for a unified action.

Temporary random networks offer a very effective way of bringing people together. Participants become more integrated into a larger community, while still retaining the specificity and individual character pertaining to them and to their smaller subgroups.

Dissemination Strategies and Sub-communities

The smaller sub-communities within the network play an important role in disseminating information. What are they, exactly? Blondel et al [10] propose a very efficient algorithm to detect communities within networks. At the outset each node belongs to its own community, in other words, there are as many communities as there are nodes. Everyone is on their own. The modularity (or the density of links within the communities compared to links between the communities) of such network is zero. Then as we start pairing the neighboring nodes, we consider they belong to the same community when the gain in modularity is the highest. In other words, when the node connects to another node and both of them are connected to some other same node, the three form a community, because the links between them are more dense than between their community and its surroundings. This also goes in line with the research in node similarity, where it was proposed that the nodes are similar when their neighbors share the same neighbors (in other words, they have a similar structural position in the network and belong to the same community). [11]



There is a lot of interesting research done on disease propagation in networks and vaccination strategies. Ball et al [12] propose a "global" vaccination strategy where a proportion of susceptibles in every group is vaccinated, rather than proportion of the population. This way, the probability of the global spread is reduced. In our context this means that if we want to spread

information to the whole community, it's not enough to randomly address a few nodes or even hubs. It is quite likely then that this information will spread within their sub-communities, but not throughout the network as a whole. Instead, any dissemination strategies should be directed at groups as a whole. Once several groups in the network pick up on a certain piece of information, they will transfer it to the other groups along the "supernetwork" that connects them and forms the giant component. Otherwise, if the information is only communicated to a part of the group, it will first have to spread among the group in order to make an impact on the whole network. This might happen so slow that by the time the information reaches the rest of the network, the original group is already "immune".

Social networks are often assortative: the nodes in the network that have many connections tend to be connected to other nodes with many connections [13]. This has several implications. First, it was shown that assortative networks spread disease (or information) more easily than disassortative ones. The explanation is that the most connected hubs form a subnetwork with a higher average degree than the network as a whole. Therefore, information travels easily and quickly from hub to hub and then gets disseminated within the cluster the hub belongs to. Second, assortative networks, just like scale-free networks, are robust to random node removal and can function even when a large fraction of the members quit.

Third, assortative networks were found to synchronize not as well as disassortative ones [14]. "[Synchronization] occurs when many interacting agents, characterized by a specific dynamical behavior, stabilize in a common evolution over time. This process can be modeled as a network of coupled oscillators (dynamical systems) with a general complex topology, synchronizing over a common evolution." [14] An example of diassortative networks that synchronize well is a neural network, a web of internet routers, protein interactions, and food chains. [13]

Random networks are disassortative too. Within the network we used as an example there are smaller disassortative ones. As they can mobilize better, they might be a good starting point to create a change within the network.

The implications for social networks is that although it's quite easy to propagate information through them and they are robust to a random node removal, it's harder to bring such networks to synchronized movement, mobilization, or common action and evolution. In order to do that, the hubs would need to start connecting to the periphery, or the periphery itself would need to start creating its own hubs. Again, like in the example above with scale-free networks, more

randomness brings the possibility of synchronization and mobilized action. However, too much randomness also makes it possible to spread malicious information too fast. Therefore, in order for the network to develop in a way that is innovative and sustainable at the same time, there should be both inclusive and exclusive activity happening simultaneously, inclusive exclusivity. Both the hubs and the periphery should be seeking new random connections, while disposing the existing ones in a more or less random manner.

The Quality of Connection and the Truth

So far it has been assumed that the quality of connections that every node has is the same. However, some nodes interact more often than others and the intensity of these interactions also differs across the network. The easiest way to measure the quality of connection is to estimate the amount of information that passes through it during a certain period of time. In real life this could be done, for example, by counting the number of messages that people exchange. However, in cybernetics information is usually measured in terms of entropy (or disorder) that it reduces [15], entropy being the logarithm of the variety of the system's states (given that the probability for these states to occur is equal from the observer's point of view). [16] In other words, it's not about the number of messages, but the ability of information to reduce uncertainty and make a difference that really counts.

Countless Twitter and Facebook updates ("London", "New York", "Paris") circulating between populations of strangers related through categorical semantic tags ("friend", "met", "neighbor") is the perfect illustration of information that is generating uncertainty, rather than reducing it. What could have once made a difference is now disjointed into small bits and pieces of utterances constantly retranslated in order to keep the network alive. The freedom is substituted by mobility (as long as one has a proper "ID" or password to pass numerous checkpoints), personal ties are discarded in favor of endlessly proliferating one's image, and neutrality (or simple indifference) is a common modus operandi. But what kind of network is it? What is the reason for it to be connected at all? Is it possible to find a way to come together that does not have anything to do with fulfilling the social (and now also virtual) obligations?

Each time information passes through a network it will proliferate the principle which the network is founded upon. When "networking" is used in a business context, it's about subjugating the power of interconnectedness to money, this lowest common denominator of social relationships. Social networks offer a more interesting trade-off: stay "connected" and "open" in exchange for centralizing all your information and creating a profile page. The main principle there is vanity, because dealing with `people as images and IDs rather than complex human beings makes it much easier to sell stuff.

Networks need a different organizing principle, the connecting force that will be closer to some internal truth (as subjective as it may sound), rather than certain agendas, global business and security plans.

"An encounter, a discovery, a vast wave of strikes, an earthquake: every event produces truth by changing our way of being in the world. Conversely, any observation that leaves us indifferent, doesn't affect us, doesn't commit us to anything, no longer deserves the name truth. There's a truth beneath every gesture, every practice, every relationship, and every situation. We usually just avoid it, manage it, which produces the madness of so many in our era. In reality, everything involves everything else. The feeling that one is living a lie is still a truth. It is a matter of not letting it go, of starting from there. A truth isn't a view on the world but what binds us to it in an irreducible way. A truth isn't something we hold but something that carries

us. It makes and unmakes me, constitutes and undoes me as an individual; it distances me from many and brings me closer to those who also experience it. An isolated being who holds fast to a truth will inevitably meet others like her. " [17]

The only shared "truth" that is available to individuals within a network is their shared desire to be together. When a network is bound by the "truth" (whatever it may be), it is bound by affect, not reason. It doesn't need an agenda, a manifesto, a leader. The participants rejoice in this being together as distinct individuals who come together because they share a common event, even if it's just a momentary fascination with each other.

"Communes come into being when people find each other, get on with each other, and decide on a common path. The commune is perhaps what gets decided at the very moment when we would normally part ways. It's the joy of an encounter that survives its expected end. It's what makes us say "we," and makes that an event.
[...]
Communes that would not define themselves – as collectives tend to do – by what's inside and what's outside them, but by the density of the ties at their core. Not by their membership, but by the spirit that animates them." [17]

What is it if not the "occasion of experience", a temporary act, an event of coming together based on affective relationship more than anything else.

"'With' implies proximity and distance, precisely the distance of the impossibility to come together in a common being. That is for me the core of the question of community; community doesn't have a common being, a common substance, but consists in being–in–common, from the starting point it's a sharing, but sharing what? Sharing nothing, sharing the space between." [18]

Dmitry Paranyushkin is an artist, researcher, curator, and media entrepreneur working primarily with live performance, image, and text. He created an online network exploration and editing tool ThisIsLike.Com that proposes to conceive of knowledge as a network of associations, rather than fixed categories and hierarchies.
Dmitry is interested in dysfunctional interfaces, networks, non-equilibrium stability, Belousov-Zhabotinsky reaction, and having more than two choices but less than four. Dmitry was born in Moscow in 1981 and currently is based between Berlin, France, and England. He can be contacted via his website www.deemeetree.com or by e-mail d@deemeetree.com

Resources

[1] Alfred North Whitehead. "Modes of Thought" (New York: Macmillan, 1938)

[2] Andreas Noack, "Energy Models for Graph Clustering" (Journal of Graph Algorithms and Applications, 2007)

[3] Newman, Barabasi, Watts, "The Structure and Dynamics of Networks" (Princeton University Press, 2006)

[4] Mathieu Jacomy, "ForceAtlas" algorithm - www.gephi.org

[5] L C Freeman, "A set of measures of centrality based on betweenness" (Sociometry, Volume 40, Issue 1, 2007)

[6] Albert, Jeong, Barabasi, "Error and Attack Tolerance of Complex Networks" (Letters to Nature 406, 2000)

[7] Keeling, "The Effects of Local Spatial Structure on Epidemiological Invasions" (The Royal Society, 1999)

[8] Kuperman, Abramson, "Small World Effect in an Epidemiological Model" (Physical Review Letters, 86-13, 2001)

[9] Paranyushkin, "Inclusive Exclusivity" (2010)

[10] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebre. "Fast unfolding of community hierarchies in large networks" arXiv (2008)

[11] Leicht, Holme, Newman, "Vertext Similarity in Networks" (University of Michigan, 2005)

[12] Ball et al, "Epidemics with Two Levels of Mixing" (The Annals of Applied Probability, Vol 7, No 1, 2007)

[13] Mark Newman, "Assortative Mixing in Networks" (ArXiv, 2002)

[14] M. Di Bernardo, F. Garofalo, F. Sorrentino, "Effects of Degree Correlation on the Synchronization of Networks of Oscillators" (International Journal of Bifurcation and Chaos, Vol 17, No 10, 2006)

[15] Gilbert Newton Lewis, Letter to Irving Langmuir (5 Aug 1930)

[16] W. Ross Ashby, "An Introduction to Cybernetics" (1956)

[17] Invisible Committee, "The Coming Insurrection", http://tarnac9.wordpress.com/texts/the-coming-insurrection/ (2008)

[18] Jean-Luc Nancy (discussion at EGS, 2001)