Advances in Complex Systems

SOCIAL DYNAMICS IN A LARGE-SCALE ONLINE GAME

MICHAEL SZELL[*,†] and STEFAN THURNER[*,§,¶,‡]

[*] *Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria*

[§] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

[¶] *IIASA, Schlossplatz 1, 2361 Laxenburg, Austria*

[†] *michael.szell@meduniwien.ac.at*
[‡] *stefan.thurner@meduniwien.ac.at*

Received (received date)
Revised (revised date)

Complex systems – when treated as systems accessible to natural sciences – pose tremendous requirements on data. Usually these requirements obstruct a scientific understanding of social phenomena on scientific grounds. Due to new developments in IT and collective human behavior, new dimensions of data sources are beginning to open up. Here we report on a complete data set of an entire society, consisting of over 350,000 human players of a massive multiplayer online game. All actions of all players over three years are recorded, including communication behavior and establishment of social ties. In this work we review the first steps undertaken in analyzing this vast data set, focusing on social dynamics on friend-, enemy- and communication networks. This new data-driven approach to social science allows to study socio-economic behavior of humans and human groups in specific environments with unprecedented precision. We propose two new empirical social laws which relate the network properties of link weight, overlap and betweenness centrality in a nonlinear way, and provide strong quantitative evidence for classical social balance assumptions, the weak ties hypothesis and triadic closure. In a first-time analysis of large-scale multirelational networks we discover systematic deviations between positive and negative tie networks. Exploring such virtual "social laboratories" in the light of complexity science has the potential to lead to the discovery of systemic properties of human societies, with unforeseen impact on managing human-induced crises.

Keywords: Social network analysis; Social balance; Mobility; Massive multiplayer online game; Quantitative social science.

1. Introduction

In this contribution we review some highlights of a recent series of studies of social dynamics in a large-scale, virtual human society [40, 41, 42, 43]. The work is motivated by the two fundamental problems associated with the social sciences of group dynamics since its existence: first, human collective dynamics constitutes a *complex system*, i.e. it influences its proper boundaries on which it depends. In other words human collective dynamics is context-dependent. For this reason it is not sufficient to measure only the dynamics of the constituents of the system (actions, decisions, movements, communication, etc., of humans), but it is imperative to take into account – at the same time – the surroundings (circumstances) of these constituents. This means that data requirements for complex systems are in general much higher than for simple physical systems, leading to the second fundamental problem: exactly where most needed, data in social sciences is comparably scarce and often of poor quality [26,45]. Until now, the needed quantities and quality of data on human societies was plainly impossible to obtain. Traditional methods of data collection in the social sciences such as questionnaires, polls, etc. are not only incapable of delivering the required data density, they also introduce well-known biases, i.e. the experiment influences the system [6].

Complex systems, such as human societies, consist of many locally interacting agents. Their actions are often influenced by their local surroundings as well as by global boundaries of the entire system. When these interactions are sufficiently strong and long-range, this may result in non-linear feedbacks leading to the often unexpected properties of complex systems which make them so hard to understand, predict and usually impossible to manage and control. Without a radically better understanding of collective human behavior, there is little hope to improve the handling of human-induced crises which are usually the most devastating [20].

Here we propose to follow a new approach in social sciences based on the following philosophy: from the simultaneous measurements of (i) the microscopic behavior (interactions, decisions, etc.) of humans within their surroundings and (ii) the macroscopic phenomena (systemic, collective, aggregate dynamics) emerging from these microscopic interactions it should be possible to derive a *statistical mechanics* of societies [8], i.e. a prescription of how to aggregate microscopic (individual) behavior to a systemic (societal) scale. Knowledge about this prescription might open possibilities to manage systemic dynamics, which is often large-scale and associated with tremendous costs.

As a first step in this direction we recorded practically *all* actions of *all* players taken in the virtual world of the self-developed, proprietary massive multiplayer online game (MMOG) `Pardus`. MMOGs provide a new tool for understanding human collective phenomena and social dynamics on a hitherto unthinkable scale [2, 10]. Data collected from `Pardus` at rates comparable to physical experiments, allows to conduct *complete* measurements of socially interacting humans, where subjects do not consciously realize the act of measurement. Contrary to traditional social sciences, in MMOGs the number of subjects can reach several millions, with billions of recorded actions. These actions of individual players are known in conjunction with their surroundings. This offers the unique opportunity to study a complex social system: specific outcomes of decisions can be measured, conditions under which individuals take decisions can in principle be controlled. In this respect social science is on the verge of becoming a fully experimental science [26] which should increasingly become capable of making a great number of repeatable and eventually falsifiable statements about collective human behavior. It is not obvious *a priori* that a population of online players is a representative sample of real-world societies [46]. Also, behavior in an online environment may be influenced by the anonymity of the users and does not necessarily reflect behavior of humans in offline environments where reputation effects often play an important role. For example, it is known that in certain online environments a fraction of users frequently act out different personalities or even genders due to the absence of real-life social repercussions [22]. However, several recent studies are providing evidence that statistical differences of real-world communities and game-societies are often marginal [23, 24]. One reason for a generally good overlap between online and offline personality of users in MMOGs might be the substantial investment of time and emotions into their online characters and their online reputation [42].

2. The Game

`Pardus` (`http://www.pardus.at`) is a browser-based MMOG in a science-fiction setting, open to the public and played since September 2004. Browser-based MMOGs are characterized by a substantial number of users playing together in the same virtual environment connected through an internet browser [5, 9]. Players live and act within a virtual, open-ended and persistent futuristic universe, making up their own goals. They claim territories, engage in economic activities, self-organize within groups, decide to go to war, etc., completely on their own accounts. Typically players participate from several weeks up to years [42].

In this virtual world every player owns a spacecraft with cargo capacity, which allows to roam the universe, to produce and trade commodities, socialize, etc., "to gain wealth and fame in space" (`http://www.pardus.at/index.php?section=about`). Main driving forces in `Pardus` are the possibility to trade and to engage in social life such as friendship, cooperation or competition. There are a number of well-used ways to publicly display one's "status" within the virtual society: accumulation of (expensive) status symbols, medals of honor for war efforts, altruistic behavior, etc. Presently more than 350,000 players have registered at `Pardus`, on a daily basis it is actively played by $\approx$ 12,000. Of the three available, independent game universes, here we focus on *Artemis* which was opened in June 2007 and accommodates $\approx$ 6,500 active players. Data is available for up to 1,238 consecutive days.

2.1. *Types of social interactions*

Players can anonymously mark others as *friends* (F) or *enemies* (E), for any reason. The marked players are added to the marker's personal *friends* or *enemies list.* Additionally, every player has a personal *friend of* and *enemy of* list, displaying all players who have marked them as friend or enemy, respectively. Private Messages (PM) are the prevalent form of communication (C) within the game. It is a system

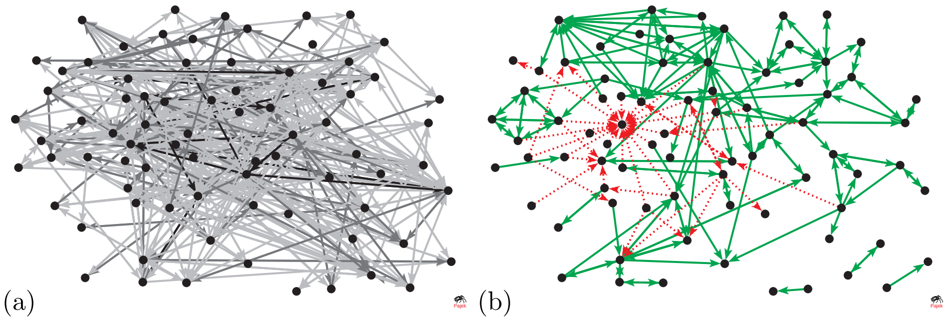


Fig. 1. (a) Accumulated PM communications over 445 days between 78 individuals. Link colors of light gray, gray, and black correspond to 1–10, 11–100 and 101–1000 PMs sent, respectively. (b) Friend (green, solid) and enemy (red, dashed) relations on one day between the same individuals. See `http://www.youtube.com/user/complexsystemsvienna` for animated time evolutions of these networks. (following [42])

similar to email – a PM is only seen by sender and receiver. There are three more types of social interaction: Trade (T), attack (A), and placing a bounty on another player (B). The removal of friendship (D) and enmity (X) links are further actions. The action types C, T, F, amd X can be associated with *positive* actions, while A, B, D, and E are *negative* actions.

All relations can be displayed as networks; see Fig. 1 for a visualization of PMs and friend/enemy relations between a small number of players.

2.2. *Types of social networks of the players*

Networks are represented as directed graphs [44]. Nodes represent players, links indicate friendships, enemies, communication, trade, attacks, etc., in the respective networks. Here we focus on three networks.

2.2.1. *Communication networks*

A set of networks is extracted by considering all PM communications on a weekly timescale. A weighted link pointing from node i to node j is placed if player i has sent at least one PM to player j within a given week. Weights correspond to the total number of PMs sent within this week. Figure 1 (a) illustrates a subgraph of PM networks.

2.2.2. *Networks of friends and enemies*

Friend and enemy markings define the following networks: a link is placed from node i to j if player i has marked player j as friend/enemy. Note that friend/enemy markings exist until they are removed by players (or as long as the players exist), while PM networks are constructed through an accumulating process. Figure 1 (b) illustrates a subgraph of a friend/enemy network.

3. Statistical Results on Behavioral Streams

In a first attempt to understand the nature of the constituents of the system, the players, we focus on the statistical properties of how players are performing actions and reactions without considering the topology of the networks spanned by these actions [43]. Here the data set includes all actions performed by all 34,055 players active in the first 1,238 days of the game universe Artemis. For the analysis, we discard players with a history of less than 1,000 actions, leaving a set of 1,758 players and their 'behavioral codes'. The action stream of a player i is defined as his time-ordered stream of N actions in his 'life', $A^i = \{a_n | n = 1, \dots, N\}$, where the a_n are any of the 8 previously defined action types. Similarly we denote $R^i = \{r_n | n = 1, \dots, M\}$ as the ordered set of received actions of a player i. The chronologically combined sequence C^i is the set of player i's actions and received-actions, having length $N+M$. We do not consider actual times between consecutive actions but only the time ordering – we work in 'action-time'.

We first analyze the transition probability matrix formed by all actions and received actions in all C^i sequences. By $p(Y|Z)$ we denote that an action of type Y directly follows an action of type Z. The ratio $\frac{p(Y|Z)}{p(Y)}$ additionally accounts for the different frequencies of action types and - if it deviates from 1 - indicates if there are correlations between subsequent actions or received actions. Figure 2 (a) and (b) shows the transition matrix of all actions and received actions (the latter marked by the subscript r), and the ratios $\frac{p(Y|Z)}{p(Y)}$ for positive and negative actions or received actions, respectively. From these measured values we find the following results: (i) The diagonal in Fig.2 (a) shows that most actions are highly repetitive, and that communication displays a distinct back-and-forth nature (larger values for $C \to C_r$ and $C_r \to C$ than for $C \to C$ and $C_r \to C_r$). (ii) The probability to perform a good action is significantly higher if previously a good action has been received, and vice versa. (iii) Negative behavior, especially attack, is highly persistent. Further analysis shows that the probability to perform a negative action is significantly higher if previously a negative action has been received, compared to the case where a positive action has been received.

Applying anomalous fluctuation techniques as previously developed for the study of e.g. DNA sequences and economic timeseries [39] on the 'world-lines' of action streams (a world line goes up (down) if a good (bad) action was performed) quantifies the persistence or anti-persistence of action types and reveals further findings [43]. For example, the vast majority of players are 'good', and the few 'bad' players tend to be short-lived and dominant, i.e. they quit the game after fewer actions than good players and they perform significantly more actions than they receive. We interpret these findings as empirical evidence for self-organization towards reciprocal, good conduct within a human society.

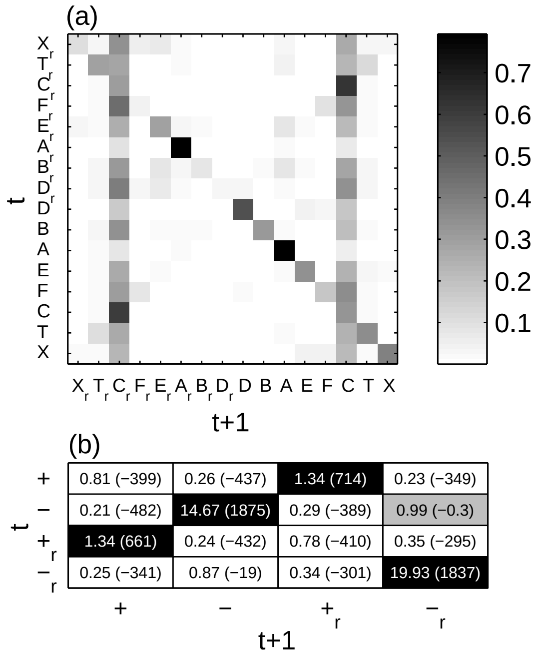


Fig. 2. (a) Transition probabilities $p(Y|Z)$ of all consecutive $(t \to t+1)$ actions and received actions in all C^i sequences of actions and received actions (the latter marked by the subscript r). High values in the diagonal show that actions tend to be repetitive. Communication is the most commonly performed (and received) action, and it displays a distinct back-and-forth nature (larger values for $C \to C_r$ and $C_r \to C$ than for $C \to C$ and $C_r \to C_r$). (b) The ratios $\frac{p(Y|Z)}{p(Y)}$ for positive and negative actions or received actions, additionally accounting for the different frequencies of action types. Deviations from 1 indicate correlations between subsequent actions or received actions, Z-scores (number of standard deviations) are shown in brackets. Positive correlations are shown on black backgrounds: The probability to perform a positive action is significantly higher if previously a positive action has been received, and vice versa. Negative behavior is highly persistent. (following [43])

4. Subdiffusive Mobility and Socio-economic Borders

Diffusion processes in complex systems often do not follow Gaussian statistics [30]. In particular, instead of Brownian motion, i.e. a linear time dependence of the mean square displacement (MSD)[a], $\sigma^2(t) \sim t$, often anomalous diffusion is observed, typically in the form of a non-linear, power law growth of the mean square displacement

[a]The MSD is a standard measure in physics, measuring the displacement of mobile objects (such as particles) over time. It is defined as $\sigma^2(t) = \langle (r(T+t) - r(T))^2 \rangle$, where $r(T)$ and $r(T+t)$ are the locations a player occupies at times T and $T+t$ respectively, and where $(r(T+t) - r(T))$ denotes the distance between the two locations. The average $\langle \cdot \rangle$ is performed over all windows of size t, with their left boundaries going from T=0 to T=1,000-t, and over all players in the considered data set.

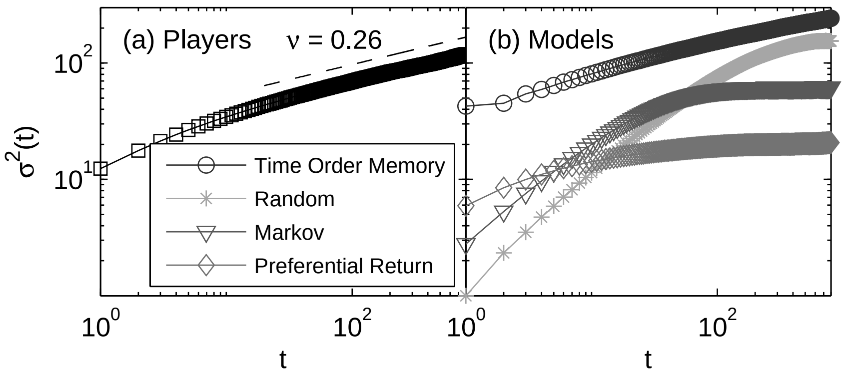


Fig. 3. Mean square displacement (MSD) of (a) players and (b) models, the time scale is given in days. The MSD of players follows a power relation $\sigma^2(t) \sim t^\nu$ with a subdiffusive exponent $\nu \approx 0.26$. This functional form is well reproduced by a 'TOM model' (Time Order Memory) which takes into account the measured long-range correlations of previously visited locations. Model curves are shifted vertically for visual clarity. (following [41])

$\sigma^2(t) \sim t^\nu, \nu \neq 1$. This functional form is strongly related to the breakdown of the classical central limit theorem due to heavy tail distributions or long-range correlations.

A particular kind of diffusion process in social systems is the mobility of humans. Understanding the statistical patterns of human mobility is a considerable challenge with important applications to traffic management [19], epidemiology [3], or information spreading [32]. Large amounts of data on various human activities, most importantly mobile phone records [14], have recently been used as a proxy for human movements. These studies have provided insights into several aspects of human mobility, uncovering distinct features such as scaling laws and anomalous diffusion [38].

In the reviewed mobility study [41], we analyze the raw data of the daily positions of the 1,458 most active players in their game universe over 1,000 days. We find that players move in a highly subdiffusive fashion, showing an exponent of $\nu \approx 0.26 \ll 1$ in the MSD, see Fig. 3 (a). The subdiffusion stems from long-range correlations in the return to previously visited locations. We use this insight to construct the 'TOM model' (Time Order Memory) which reproduces well the slope, see Fig. 3 (b). The TOM model incorporates the distributions of first return times (power law), together with the measured distribution of waiting times (power law) and the distribution of jump distances (exponential). A simple Markov model with perfect information on first order transition probabilities, as well as the previously suggested preferential return model [38] fail to reproduce the measured MSD because they overemphasize preferences of locations visited long ago.

Further, we suggest a technique to recover socio-economic regions in the game

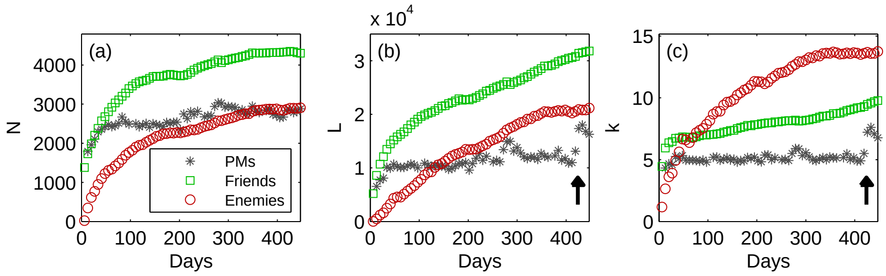


Fig. 4. Evolving network properties: (a) number of nodes N, (b) number of (directed) links L, (c) average degree $\bar{k}$. Arrows mark the outbreak of an in-game war at day 422. (following [42])

universe almost perfectly, by applying community detection methods to the raw movement data. This method works because - as we are able to quantify - players have a significant tendency to avoid crossing borders of 'countries' [41]. This information is of importance for understanding the role of political or socio-economic borders on the migration of humans, where the presence of such borders can have a strong influence on mobility [36].

5. Socio-dynamic Results using Networks

We measure the time evolution of basic network properties such as: number of nodes N and directed links L, average degree $\bar{k}$, Fig. 4, for technical details see [42]. Cumulative distributions of in- and out-degrees of the networks are depicted in Figs. 5 (a) and (b). Note how out-degrees (i.e. the amount of friends / enemies / communication partners) are limited by $k^{\mathrm{out}} \approx 150$. This well-known "Dunbar number" [12] is assumed to be a natural limit of group sizes of humans and primates due to limited cognitive capacities.

5.1. *The network densification effect*

The networks studied in [42] confirm the observations of growing average degrees, Fig. 4 (c), and shrinking diameters (not shown). This means that over time, people in a social network become *closer*, in the sense that e.g. circulating information has to be passed on by fewer and fewer people for reaching a destination in the network. This *network densification* [29] or *accelerated growth* [11] is in contrast to simple network growth models such as preferential attachment (PA) [4]. The model of PA assumes that nodes which link to a network for the first time preferably attach to "popular" nodes [4]. Growing average degrees were also observed in recent dynamical network studies [21, 29, 34, 37].

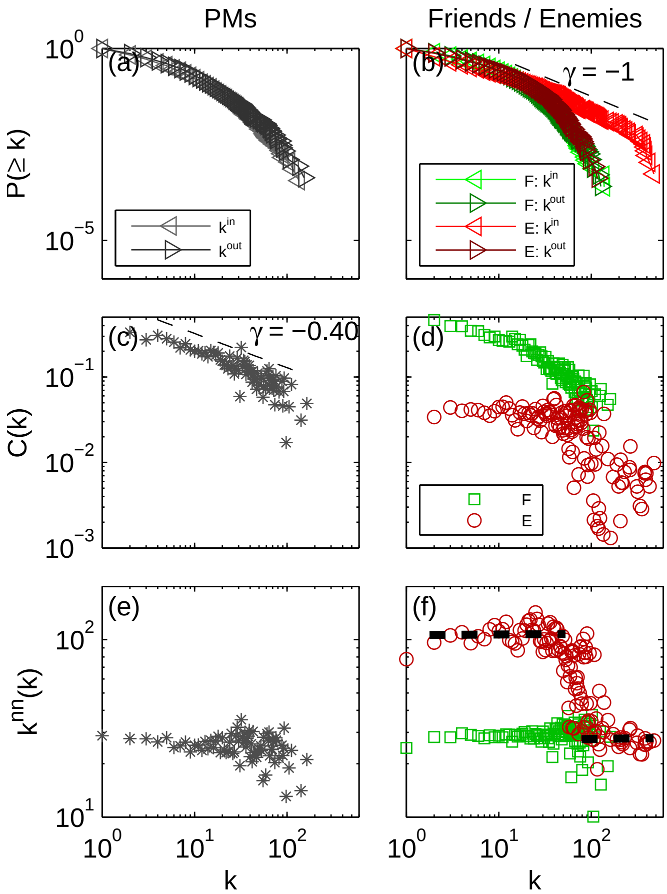


Fig. 5. Cumulative degree distribution of (a) PM, (b) friend and enemy networks; clustering coefficient C as a function of degree for the (c) PM, (d) friend and enemy networks; nearest neighbor degree k^{nn} versus degree of the (e) PM, (f) friend and enemy networks. Thick dashed lines mark two classes of enemies. Networks are shown at day 445. (following [42])

5.2. *Asymmetry of friend and enemy networks*

If a network follows a PA model three facts should be observable. (i) the linking probability $P(k)$ is proportional to k^{α}, with $\alpha = 1$; (ii) the degree distribution follows a power-law $p(k) \sim k^{-\gamma}$, and (iii) the clustering coefficient c_i is uncorrelated with the degree k_i.

None of these facts are present in the friend networks, Figs. 5 (b) and (d), and 6.

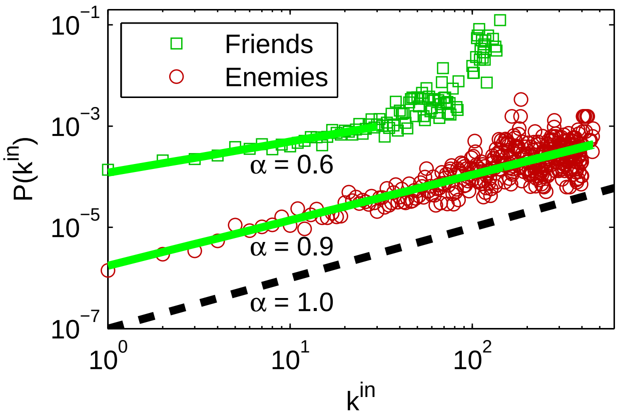


Fig. 6. Empirical probability $P(k^{\mathrm{in}})$ for newcomers connecting to existing nodes with given in-degree k^{in}. The black line depicts slope $\alpha = 1$ and indicates the assumption needed in the PA model. Green lines denote least squares fits. Values for enemies are vertically displaced for better visibility. (following [42])

The linking probability exponent of $\alpha = 0.6$ deviates significantly from $\alpha = 1$.[b] The degree distribution is hardly a power-law, and the clustering coefficient depends on k, Fig. 5. For enemy networks the situation is different. The linking probability exponent is $\alpha \approx 0.9$, the distribution of in-degrees is closer to an approximate power-law with exponent $\gamma \approx 1$, being consistent with PA. Asymmetries between network formation processes of friend and enmity networks are explainable by e.g. social penetration theory [25] or social balance theory, see below.

5.3. *Two categories of enemies*

Identifying negative social tie mechanics may be much more important for gaining insight in social group dynamics than identifying mechanics of positive social ties [25]. Our measurements provide first steps toward this direction.

In the plot of average neighbor degree k^{nn} versus degree k one observes two classes, see dashed lines in Fig. 5 (f). The first class contains players with low degrees ($k < 50$), all having an average neighbor degree of $k^{\mathrm{nn}} \approx 100$. The second class has players with a degree of $k > 100$ and an average neighbor degree of $k^{\mathrm{nn}} \approx 28$. There is a sharp transition between these classes. These observations suggest two distinct mechanics of enemy formation at work:

(1) *Vendetta – private enemies:* A player who directly experiences a negative action by another player is likely to *immediately* react by marking this agitator as

[b]The same exponent of $\alpha = 0.6$ has been measured for LinkedIn [27], a social networking site for professional contacts.

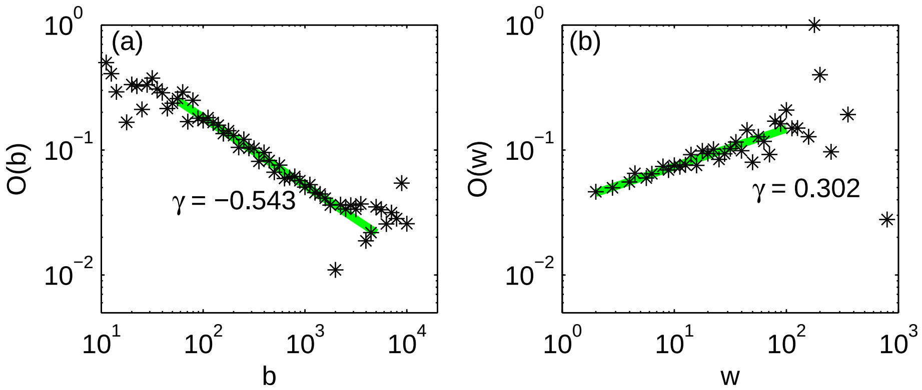


Fig. 7. (a) Overlap versus betweenness and (b) overlap versus weight in a PM network. Stars denote log-binned averages of overlap values of the links, lines are least-squares fits. (following [42])

enemy. In this scenario, a dyadic vendetta without involvement of other players takes place. Due to the absence of long-lasting repercussions on real-life reputation, negative dyadic behavior in online environments might be artificially inflated. For example, studies on the collaborative online encyclopedia *Wikipedia* have shown extended negative phenomena such as vandalism and so-called *edit wars*, i.e. repeated edits and reverts between pairs of editors or pairs of editor groups representing opposing opinions [47].

(2) *Volksfeind – public enemies:* Some players display a destructive personality [9], taking pleasure in destroying other players' work. Anonymity of the internet facilitates this behavior due to lack of real-world social repercussions. For this reason these few troublemakers tend to cause a lot of misdoings to a big number of players. After such an individual is identified by the community, she may receive pre-emptive enemy markings, either by friends of offended friends or by otherwise non-involved players. This destructive behavior and the indirect marking mechanism leads to the emergence of a few "public enemies", i.e. players marked as enemy by a huge number of others.

As a counterpoint to the somewhat vague situation of how far negative behavior might deviate in online environments from negative behavior in offline environments, in the following sections we provide empirical evidence for several well-known hypotheses in social science concerning positive and communication behavior.

5.4. *Confirmation of the weak ties hypothesis*

A long-standing proposition in sociology, the *weak ties hypothesis*, builds upon the assumption that "the degree of overlap of two individual's friendship networks varies directly with the strength of their tie to one another" [15]. Weak ties (e.g. casual

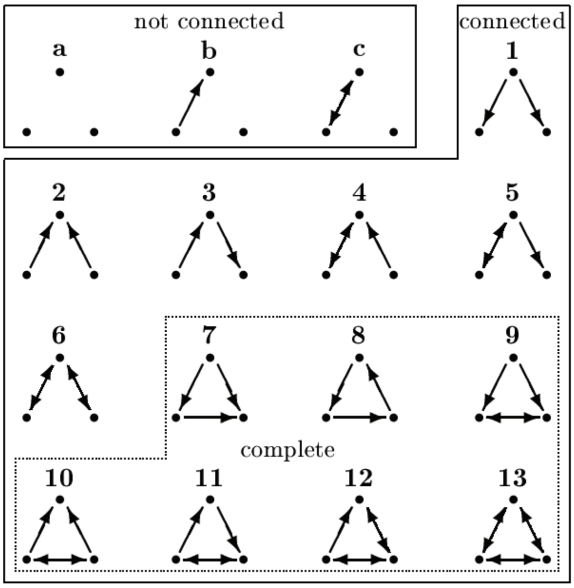


Fig. 8. The 16 classes of triads and their ids. The transition id 6 → id 13 represents one example for triadic closure. (following [42])

acquaintanceships) are assumed to be strong in the sense that they weakly link communities characterized by strong ties (standing for e.g. good friendships). Strong ties correspond to redundant connections *within* communities. As an intuitive notion of *strength* of an interpersonal tie, Granovetter mentions "the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie". These hypotheses can be tested by examining the measured function of overlap [33] versus PM weight, $O(w)$, as well as overlap versus link betweenness centrality [13], $O(b)$. For a definition see [42]. Intuitively, the former should be increasing, the latter decreasing. In mobile phone call networks [33] exactly this behavior was reported. More quantitatively, our data set suggests an approximate cube root law and an inverse square root law relating the fundamental network parameters betweenness and link-weight to overlap [42], see Fig. 7:

$$O(w) \sim \sqrt[3]{w} \qquad \text{and} \qquad O(b) \sim \frac{1}{\sqrt{b}} \tag{1}$$

Further tests on other large-scale communication networks, are needed to determine if these data-derived "social laws", eq. (1), describe universal patterns of human communication.

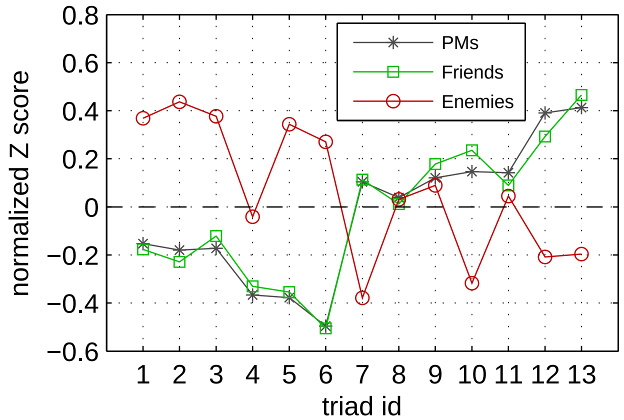


Fig. 9. Triad significance profiles of the three networks (day 445). (following [42])

5.5. *Confirmation of triadic closure*

Triads are 3-node subgraphs. In directed graphs there exist 16 classes [17], Fig. 8 (a). The *triad significance profile* (TSP) of a given network reveals which of these classes are over (under) represented with respect to a random null-model [31]. The Z-score quantifies the degree of overrepresentation of a triad class: positive (negative) Z-score means over (under) representation. The TSP can be used to test the *triadic closure hypothesis* [15], stating that within a social network of positive ties (e.g. friendships) triad class 6 should have the smallest Z-score whereas class 13 should have the highest. Class 6 is a "frustrated" state where one person has two friends, but these friends do not know each other; class 13 is the state where all three are friends. The phenomenon of *triadic closure* [35] states that individuals are driven to fill the "hole" in triad class 6. More generally, if we focus on completeness, i.e. whether a hole exists or not, we expect negative Z-score of the incomplete triad classes 1–6 and positive Z-score of the complete triad classes 7–13. We find excellent empirical agreement with Granovetter's prediction for friend and PM networks: triad 6 has minimal, class 13 maximal Z-score, Fig. 9. Our findings further confirm the TSPs of the superfamily of social and hyperlink networks found in [31] and other social networks [16]. So-far triad dynamics on networks of *negative ties* has not been measured on large scales. Following similar social balance arguments we expect reversed roles of completeness: instead of the *absence* of a completing third link, its *presence* should cause frustration. Hence in enemy networks triad classes 1–6 should be overrepresented, triad classes 7–13 underrepresented. We observe confirmation of this "mirrored" hypothesis: most Z-scores in enemy networks have opposite signs of those in friend networks; the deviations (ids 4, 9, 11) reveal interesting peculiarities.

Explicit evidence for triadic closure can be provided by directly counting transitions between triads. Entries in the 13×13 matrix $\mathbf{K}$ give the number of transitions

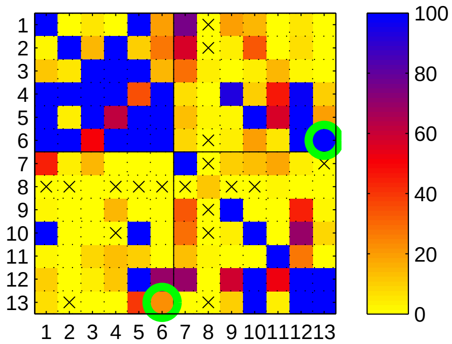


Fig. 10. Matrix **K** of empirical 50-day transition counts between triad classes in friend networks. Blue squares mark average transition counts ≥ 100. Crosses mark transitions which never occured. Circles point out the asymmetry of transitions between triad classes 6 and 13. (following [42])

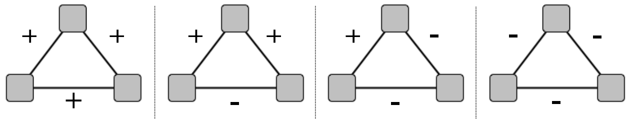
from one triad to another within a time interval of 50 days, here in the friend networks, Fig. 10. According to the hypothesis of triadic closure, $K_{6,13}$ should contain high values,while the other way $K_{13,6}$ is suppressed. We find that this is the case, $K_{6,13} = 306 > 23 = K_{13,6}$. In general, we observe incomplete $\rightarrow$ complete transitions (upper right sector in **K**) between connected triad classes than vice versa (lower left in **K**). Transition rates could be essential for model builders [1]. Due to lack of data, underlying parameters for agent based models so-far could only be assumed.

5.6. *Multiplexity and confirmation of social balance*

In Ref. [40] we extend the analysis of three types of networks (PMs, friends, enemies) with three additional network types. In total we study six network types, three having a positive connotation (communication, friendship, trade), and three with a negative connotation (enmity, attack, bounty). The observation of a heavy tail in the degree distribution of the enmity network, see Fig. 5 (b), is extended and systematic deviations between positive and negative tie networks can be found: In general, negative [positive] tie networks display [no] heavy tails, much lower [higher] reciprocity (fraction of reciprocated links), and a lower [higher] clustering coefficient C.

We proceed to study the social system following a multiplex viewpoint [44]. In this approach, the set of all networks defines the *Multiplex network*, in which the nodes (individuals) can be connected by different types of links [44]. This evolving and highly non-trivial object can provide essential insights into the organization principles of a society, and is useful for testing further sociological hypotheses.

In the following we focus only on the multiplex sub-network made up of friend and enemy relations and assign a +(-) sign to a friendship (enmity) link. Triads



	B	U	B	U
N_Δ	26,329	4,428	39,519	8,032
N_Δ^{rand}	10,608	30,145	28,545	9,009
z	71	-112	47	-5

Fig. 11. The four different types of complete triads in the signed multiplex network of friendships and enmities are $+++$, $++-$, $+--$, $---$. B and U stand for balanced and unbalanced triads, according to social balance theory. The N_Δ values show the observed counts, N_Δ^{rand} are the expected numbers in the null model. Balanced triads are indeed overrepresented (Z-scores 71 and 47), and the unbalanced triad $++-$ highly underrepresented. The $---$ triad is much less underrepresented (Z-score -5) than $++-$ (Z-Score -112). (following [40])

within this signed network are positive if the product of its links is positive ('the friend of my friend is my friend' or 'the enemy of my enemy is my friend') and negative otherwise ('the friend of my enemy is my friend' or 'the enemy of my enemy is my enemy'). We perform a large-scale test of social balance, a sociological hypothesis which goes back to the 1940's [18] and claims that positive triads are 'balanced' while negative triads are 'unbalanced' [7].

For testing social balance, we count all complete triads (being either of type $+++$, $++-$, $+--$ or $---$) in the signed network, N_Δ, and compare these numbers to the expected number N_Δ^{rand} of triads in the null model which fixes the topology but randomly reshuffles the signs of the links. Figure 11 shows the values and corresponding Z-scores in the last row, measuring the significance of over- or underrepresentations. Indeed we measure highly overrepresented positive triads and a very significant underrepresentation of the negative $++-$ triad. The negative triad $---$ ('enemy of my enemy is my enemy') is much less underrepresented.[c] Similar results have been found in another study on large-scale signed networks in social media [28].

6. Conclusion

This review of [40, 41, 42, 43] demonstrates first attempts to transform traditionally non-natural sciences into sciences which allow for quantifiable and falsifiable pre-

[c]Note that the random network of the null model is only random in the sense that signs of links are randomized. This is different to a random network where the topology is reshuffled. In this latter case the negative triads $---$ are slightly overrepresented, which can be inferred from the clustering coefficient [42].

dictions. To sketch the enormous potential these new kind of scientific endeavors yield for the future, we showed several hitherto inaccessible facts on large-scale social networks, confirmed two longstanding social balance hypotheses, and suggest two predictions on "social laws", eq. (1), relating local (overlap, weight) and global (betweenness centrality) network measures in a quantitative manner.

The key ingredient in these attempts is to overcome the tremendous data requirements needed for an experimental approach to complex systems. There is little doubt that data-driven, quantitative methods of measuring and analyzing collective human behavior in large-scale "social laboratories" will greatly accelerate the progress of social science. Given joint efforts of scientists and game designers, virtual worlds such as `Pardus` which attract hundreds of thousands of players, are relatively easy to set up for specific research questions. Basically any study of aggregate human phenomena can be carried out: from behavioral economics, origin of cooperation and ethical behavior to the dynamics of conflict, war and terrorism. New insights may lead to unforeseen impacts on managing human-induced crises. It remains an open question to which extent social behavior in real society deviates from behavior of humans in virtual societies. Only continued efforts on high-frequency and large-scale studies will eventually show which of the above findings can be referred to as "universal".

Acknowledgments

We thank Werner Bayer for compiling `Pardus` data. Supported in part by Austrian Science Fund FWF P19132 and FWF P23378, by EU FP7 project INSITE, and by the European Cooperation in Science and Technology Action, COST MP0801. All data used in this study is fully anonymized; the authors have the written consent to publish from the legal department of the Medical University of Vienna.

References

[1] Antal, T., Krapivsky, P., and Redner, S., Social balance on networks: The dynamics of friendship and enmity, *Physica D: Nonlinear Phenomena* **224** (2006) 130–136.
[2] Bainbridge, W., The scientific research potential of virtual worlds, *Science* **317** (2007) 472.
[3] Balcan, D., Colizza, V., Goncalves, B., Hu, H., Ramasco, J., and Vespignani, A., Multiscale mobility networks and the spatial spreading of infectious diseases, *Proceedings of the National Academy of Sciences* **106** (2009) 21484–21489.
[4] Barabási, A. and Albert, R., Emergence of scaling in random networks, *Science* **286** (1999) 509.
[5] Bartle, R., *Designing virtual worlds* (New Riders Games, 2004).
[6] Carrington, P., Scott, J., and Wasserman, S., *Models and methods in social network analysis* (Cambridge University Press, 2005).
[7] Cartwright, D. and Harary, F., Structural balance: a generalization of Heider's theory, *Psychological Review* **63** (1956) 277–93.
[8] Castellano, C., Fortunato, S., and Loreto, V., Statistical physics of social dynamics, *Reviews of Modern Physics* **81** (2009) 591–646.

[9] Castronova, E., *Synthetic Worlds: The Business and Culture of Online Games* (University of Chicago Press, Chicago, 2005).
[10] Castronova, E., On the research value of large games, *Games and Culture* **1** (2006) 163–186.
[11] Dorogovtsev, S. and Mendes, J., Effect of the accelerating growth of communications networks on their structure, *Physical Review E* **63** (2001) 25101.
[12] Dunbar, R., Co-evolution of neocortex size, group size and language in humans, *Behavioral and Brain Sciences* **16** (1993) 681–735.
[13] Freeman, L., A set of measures of centrality based on betweenness, *Sociometry* **40** (1977) 35–41.
[14] González, M., Hidalgo, C., and Barabási, A., Understanding individual human mobility patterns, *Nature* **453** (2008) 779–782.
[15] Granovetter, M., The strength of weak ties, *American Journal of Sociology* **78** (1973) 1360–1380.
[16] Hamasaki, M., Takeda, H., Hope, T., and Nishimura, T., Network analysis of an emergent massively collaborative creation community, *Proceedings of the Third International ICWSM Conference* (2009) 222–225.
[17] Harary, F., Norman, R., and Cartwright, D., *Structural Models: An Introduction to the Theory of Directed Graphs* (John Wiley & Sons Inc, 1965).
[18] Heider, F., Social perception and phenomenal relations, *Psychological Review* **5** (1944) 358–374.
[19] Helbing, D., Traffic and related self-driven many-particle systems, *Reviews of Modern Physics* **73** (2001) 1067.
[20] Helbing, D., Managing complexity in socio-economic systems, *European Review* **17** (2009) 423–438.
[21] Holme, P., Edling, C., and Liljeros, F., Structure and time evolution of an internet dating community, *Social Networks* **26** (2004) 155–174.
[22] Hussain, Z. and Griffiths, M., Gender swapping and socializing in cyberspace: An exploratory study, *CyberPsychology & Behavior* **11** (2008) 47–53.
[23] Jiang, Z., Zhou, W., and Tan, Q., Online-offline activities and game-playing behaviors of avatars in a massive multiplayer online role-playing game, *EPL (Europhysics Letters)* **88** (2009) 48007.
[24] Johnson, N., Xu, C., Zhao, Z., Ducheneaut, N., Yee, N., Tita, G., and Hui, P., Human group formation in online guilds and offline gangs driven by a common team dynamic, *Physical Review E* **79** (2009) 66117.
[25] Labianca, G. and Brass, D., Exploring the social ledger: Negative relationships and negative asymmetry in social networks in organizations, *The Academy of Management Review* **31** (2006) 596–614.
[26] Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., *et al.*, Computational social science, *Science* **323** (2009) 721.
[27] Leskovec, J., Backstrom, L., Kumar, R., and Tomkins, A., Microscopic evolution of social networks, in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM, 2008), pp. 462–470.
[28] Leskovec, J., Huttenlocher, D., and Kleinberg, J., Signed networks in social media, *Proc. 28th CHI* (2010).
[29] Leskovec, J., Kleinberg, J., and Faloutsos, C., Graph evolution: Densification and shrinking diameters, *ACM Transactions on Knowledge Discovery from Data* **1** (2007).
[30] Metzler, R. and Klafter, J., The random walk's guide to anomalous diffusion: A fractional dynamics approach, *Physics Reports* **339** (2000) 1–77.

[31] Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., and Alon, U., Superfamilies of evolved and designed networks, *Science* **303** (2004) 1538.
[32] Miritello, G., Moro, E., and Lara, R., Dynamical strength of social ties in information spreading, *Physical Review E* **83** (2011) 045102.
[33] Onnela, J., Saramäki, J., Hyvönen, J., Szabó, G., de Menezes, M., Kaski, K., Barabási, A., and Kertész, J., Analysis of a large-scale weighted network of one-to-one human communication, *New Journal of Physics* **9** (2007) 179.
[34] Panzarasa, P., Opsahl, T., and Carley, K., Patterns and dynamics of users' behavior and interaction: Network analysis of an online community, *Journal of the American Society for Information Science and Technology* **60** (2009) 911–932.
[35] Rapoport, A., Spread of information through a population with socio-structural bias: I. assumption of transitivity, *Bulletin of Mathematical Biology* **15** (1953) 523–533.
[36] Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., and Strogatz, S., Redrawing the map of Great Britain from a network of human interactions, *PLoS One* **5** (2010) e14248.
[37] Ravasz, E., *Evolution, hierarchy and modular organization in complex networks*, Ph.D. thesis, University of Notre Dame, Indiana (2004).
[38] Song, C., Koren, T., Wang, P., and Barabási, A., Modelling the scaling properties of human mobility, *Nature Physics* **6** (2010) 818–823.
[39] Stanley, H., Afanasyev, V., Amaral, L., Buldyrev, S., Goldberger, A., Havlin, S., Leschhorn, H., Maass, P., Mantegna, R., Peng, C., *et al.*, Anomalous fluctuations in the dynamics of complex systems: From DNA and physiology to econophysics, *Physica A* **224** (1996) 302–321.
[40] Szell, M., Lambiotte, R., and Thurner, S., Multirelational organization of large-scale social networks in an online world, *Proceedings of the National Academy of Sciences* **107** (2010) 13636–13641.
[41] Szell, M., Sinatra, R., Petri, G., Thurner, S., and Latora, V., Understanding mobility in a social petri dish, *Arxiv preprint arXiv:1112.1220, in review* (2012).
[42] Szell, M. and Thurner, S., Measuring social dynamics in a massive multiplayer online game, *Social Networks* **32** (2010) 313–329.
[43] Thurner, S., Szell, M., and Sinatra, R., Emergence of good conduct, scaling and Zipf laws in human behavioral sequences in an online world, *PLoS ONE* **7** (2012) e29796.
[44] Wasserman, S. and Faust, K., *Social Network Analysis: Methods and Applications* (Cambridge University Press, 1994).
[45] Watts, D., A twenty-first century science, *Nature* **445** (2007) 489.
[46] Williams, D., Martins, N., Consalvo, M., and Ivory, J., The virtual census: Representations of gender, race and age in video games, *New Media & Society* **11** (2009) 815.
[47] Yasseri, T., Sumi, R., Rung, A., Kornai, A., and Kertész, J., Dynamics of conflicts in Wikipedia, *Arxiv preprint arXiv:1202.3643* (2012).