PLOS | ONE

RESEARCH ARTICLE

Detection of the Elite Structure in a Virtual Multiplex Social System by Means of a Generalised *K*-Core

Bernat Corominas-Murtra[1], Benedikt Fuchs[1], Stefan Thurner[1,2,3]*

1. Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A-1090, Vienna, Austria, **2.** Santa Fe Institute, 1399 Hyde Park Road, 87501, Santa Fe, New Mexico, United States of America, **3.** IIASA, Schlossplatz 1, A-2361, Laxenburg, Austria

*stefan.thurner@meduniwien.ac.at

🔓 OPEN ACCESS

Abstract

Elites are subgroups of individuals within a society that have the ability and means to influence, lead, govern, and shape societies. Members of elites are often well connected individuals, which enables them to impose their influence to many and to quickly gather, process, and spread information. Here we argue that elites are not only composed of highly connected individuals, but also of intermediaries connecting hubs to form a cohesive and structured elite-subgroup at the core of a social network. For this purpose we present a generalization of the *K*-core algorithm that allows to identify a social core that is composed of well-connected hubs together with their 'connectors'. We show the validity of the idea in the framework of a virtual world defined by a massive multiplayer online game, on which we have complete information of various social networks. Exploiting this multiplex structure, we find that the hubs of the generalised *K*-core identify those individuals that are high social performers in terms of a series of indicators that are available in the game. In addition, using a combined strategy which involves the generalised *K*-core and the recently introduced *M*-core, the elites of the different 'nations' present in the game are perfectly identified as modules of the generalised *K*-core. Interesting sudden shifts in the composition of the elite cores are observed at deep levels. We show that elite detection with the traditional *K*-core is not possible in a reliable way. The proposed method might be useful in a series of more general applications, such as community detection.

Introduction

Almost universally, across cultures and times, societies are structured in a way that a small group of individuals are in the possession of the means to influence, shape, structure, lead, and govern large proportions of entire societies. These selected minorities form the *elites*. The definition and characterization of an elite is a highly multidimensional and debated problem [1–5]. It incorporates considerations about wealth, experience, fame, influence over other individuals, role in societies, clubs, parties, etc. In any case elites can not be defined *per se*, but only within the context of a social system, which are superpositions of various time-varying social networks, so-called multiplex networks (MPN) [6–8]. These networks represent interactions between individuals as links of different types such as communication, trading, friendship, aggression, etc., see Fig. 1a. It seems natural that elites have to be defined through their location within these MPNs. Indeed, one would generally expect that members of elites are characterized by a large *connectivity* [9] in the various networks of the MPN, which enables them to
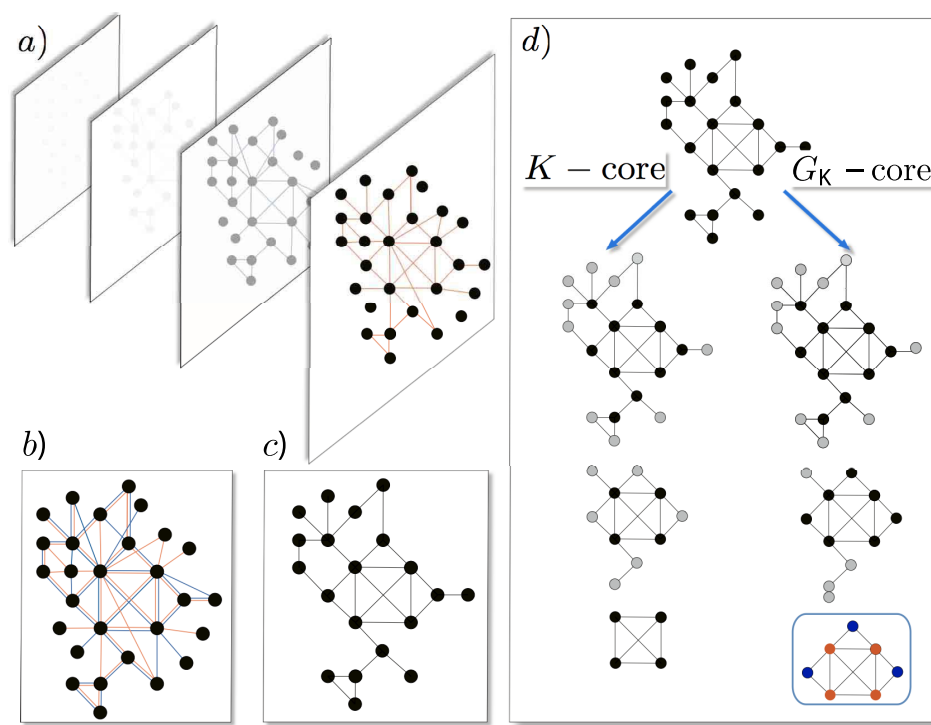
Fig. 1. Extracting the core of a Multiplex System. (a) Representation of multiplex network (MPN) composed of several layers of different relations among nodes. (b) A MPN consisting of two link-types *orange* and *blue*, and (c) its *intersection graph* obtained by keeping those links that are present on both networks. (d) Comparison of the *K*-core, left and the *generalised K*-core, right algorithms, when applied to the intersection graph: while the *K*-core iteratively removes those nodes whose degree is lower than *K*, (leading to the *K*-core), the G_K-core iteratively removes nodes whose degree is lower than *K* which are not connected to more than one node whose degree is equal or higher than *K*. We highlight the *connectors* (blue) and the *hubs* (orange). Although connectors nodes may have a low degree, they play a role in keeping the overall connectivity at deep levels of network's organization.

exert their influence on a large number of other individuals. A large connectivity, paired with a strategic position within the MPN, also allows them to collect, process, and spread information that is of relevance to them [10]. In this view elites are 'core-communities' that, to a certain extent, organise the whole topology of social interactions in a social system [9]. It is further intuitive that elites are not simply a collection of highly connected individuals, but communities of individuals densely connected (a *cohesive subgroup*) containing hubs and maybe other individuals playing functional roles within such elite structure. Moreover, relations among elite members are not incidental: they are defined at the same time at multiple levels, spanning from personal and commercial relationships to information exchanges. The cohesiveness of this group can be achieved by means of direct relations among the elite members or by means of intermediaries, individuals who, although not very connected themselves, establish and coordinate the relations between well connected elite members [11]. We refer to these intermediaries as *connectors*.

Given the above considerations, the question arises if one could identify the elite members of a given society from its MPN only by topological means. The identification of cohesive subgroups at the core of social networks has a history of decades and includes the *K*-core decomposition [12–14], the clique identification [15, 16] or the rich club analysis [17], among other general methods of cohesive subgroup identification [18, 19]. In general, these decomposition schemes are focused on the features of the organization of hubs. However, to adequately describe the organization of a social system, one might think of alternative definitions of 'core', taking into account other *functional* properties of nodes than just their degree. In the spirit of our definition of elites, connectors should be included in the definition of a core. The heart of this paper is to suggest a generalization of the *K*-core algorithm that naturally takes the 'functionality' of connectors into account, and thus allows to detect cores which are composed of hubs together with their connectors. The *generalised K*-core is obtained by an iterative method inspired both by the so-called *K*-scaffold [20, 21], and the *K*-core [12, 14]. Specifically, the *generalised K*-core (G_K-core) is the maximal induced subgraph whose nodes either have a degree larger or equal than K or *connect* two or more nodes with a degree larger or equal to K, see Fig. 1b and methods for details. We will show that G_K-cores isolate the elite communities much more reliably than the traditional *K*-cores. Moreover, as we shall see, *K*-cores and G_K-cores show substantial differences in their composition and architecture.

The quantitative exploration of structural patterns in real social systems is usually hard or even impossible due to poor data availability and due to factors that escape experimental control. Virtual societies such as those formed in Massive Multiplayer Online Games (MMOG) [22] offer an excellent opportunity to avoid these complications and allow for the first time a fully quantitative and empirical understanding of social systems under controlled conditions. Log-files of these games provide complete datasets where practically all actions and interactions of all avatars in the games are recorded. MMOGs provide a unique framework to test quantitative hypotheses and formulate entirely new questions

on social systems. Data then can provide answers at unprecedented levels of precision in the social sciences. In this paper we will use data from the MMOG society of the game 'Pardus' (http://www.pardus.at) [23], an open-ended online game with a worldwide player base which currently contains more than 420,000 people. In this game players live in a virtual, futuristic universe where they interact with other players in a multitude of ways to achieve their self-posed goals. A number of social networks can be extracted from the Pardus game, leading to the first realization of an entire MPN of a human social system. The MPN consists of the time-varying communication, friendship, trading, enmity, attack, and revenge networks. These networks are tightly related and mutually influence each other as it has been systematically explored and quantified in [7, 23–28]. Here we focus on networks representing *cooperative* interactions, namely, *friendship* (F), *communication* (C) and *Trade* (T). Our social system is therefore given by the MPN $\mathcal{M}(t) = \mathcal{M}(V, E_F \times E_C \times E_T, t)$, being E_F, E_C and E_T the sets of links defining a friendship relation, a communicative exchange or a commercial relation, respectively. To ensure the relevance of our results, we will filter the players to rule out the non-active ones. Specifically, we will build the nets over the most active players 'Artemis' universe of the game, which leads us to a set of $2000 - 2500$ players.

It is not *a priori* clear which link type of the MPN or which combination of links is most relevant for elite detection. A communication link between two individuals might signal an occasional interaction, whereas if a communication link is paired with a trade link, this might be an indication for a much stronger relation between them. For this purpose we derive four more networks, the *intersections* among levels of the MPN, see Fig. 1a,c and methods. In these networks a link exists if it is present in two or three of the MPN layers. For these intersection graphs, we formally write $\mathcal{G}_{FC} = \mathcal{G}_F \bigcap \mathcal{G}_C$, $\mathcal{G}_{FT} = \mathcal{G}_F \bigcap \mathcal{G}_T$, $\mathcal{G}_{CT} = \mathcal{G}_C \bigcap \mathcal{G}_T$ and $\mathcal{G}_{FCT} = \mathcal{G}_F \bigcap \mathcal{G}_C \bigcap \mathcal{G}_T$. The links of these networks, often called *multi-links* [29], encode strong relationships among individuals, for they connect players interacting in more than one type of relation. The strongest links in this sense are those in $\mathcal{G}_{FCT}$, a graph which we refer to as the structural *backbone* of the multiplex system. The identification of elite structures and core organization is based on the 3 networks of the MPN and their associated four intersection graphs.

The core organization of $\mathcal{G}$ will be explored explicitly by computing the sequence of G_K-cores, the so-called G_K-decomposition sequence, which amounts to a 'russian doll' decomposition of the networks,

$$... \subseteq G_K(\mathcal{G}) \subseteq G_{K-1}(\mathcal{G}) \subseteq ... \subseteq G_2(\mathcal{G}) \subseteq \mathcal{G}.$$

The behavior of this sequence of nested levels of networks (either seen in terms of the statistical properties of their graphs, or from their social composition) is essential to identify the elite organization and the elite structure of our virtual social system. When compared to the traditional *K*-core, we will see that the G_K-core provides a much more detailed picture of the nested community structures.

Data from the 'Pardus' game enables us to test and compare the quality of the identified core and to see to what extend it relates to properties that are expected for an elite. For every player we have a record of wealth, leadership role in local organizational structures, and importance in leadership as measured by a 'global leadership index'. Local organizational structures are clubs, societies and political parties, in which players organise; we know which player has a leading role in that local organization which can be president, treasurer or application master. The global leadership index is a status index that is assigned to each player (visible to all the others) which increases when special tasks (missions) are fulfilled. Such an index is an indicator of the potential influence of the player on decisions affecting the whole 'faction' it belongs to. A faction would correspond to a country in the real world. In its current state, the game extends over a universe containing three factions, which are politically independent and lead by their respective elites.

A final word of caution is needed, in relation to the significance of the data shown here. Since there is no formal/topological definition of elite in a given multiplex society, we adopted the position of showing the averages of the indicators of social relevance of the different core subgraphs we isolate. We checked the position of the topologically isolated sets of nodes within the raw rank of social performance of all players under study. However, an elite is not just a *list* of the best performers but a cohesive social structure. Therefore, rigorous indicators of statistical relevance would imply the assumption of meaningful null models. This is undoubtedly extremely interesting, but it is an issue going far beyond the scope of this paper. Instead, we adopted the position of giving relevance to our results by confronting them the the ones obtained by means of the K-core, the standard core extraction mechanism, originally designed to extract the network substructure of the most influential individuals in a given society.

Results

We extract the mentioned seven networks from the Pardus data, in the same way as described in [7, 23]. Our analysis is performed over the three networks $\mathcal{G}_F, \mathcal{G}_C$ and $\mathcal{G}_T$ obtained from the most active players in two time spans of sixty days, $t_1 = 796 - 856$ and $t_2 = 1140 - 1200$ in units of days since beginning of the game. A link between two players in the layer $\mathcal{G}_F$ exists if at least one player recognises the other as 'friend' in the whole studied period. Likewise, a link between two players in the layer $\mathcal{G}_C$ exists if at least one player has sent a message to the other in the studied time span. Finally, a link between two players in $\mathcal{G}_T$ exists if there has been at least one commercial transaction between these two players within the studied time span. The set of players that will define the set V of the MPN obtained from the period 796–856 contains 2422 players, whereas the set of players defining the MPN of the period 1140–1200 comprises 2059 players. Chosen players are those who are active in at least all three levels of the MPN during all the studied periods. The periods have been chosen using two criteria i) The periods are chosen far away enough from the starting of the game, to ensure that the social structure of

the virtual society achieved certain degree of maturity and ii) The comprised time spans do not contain 'war' periods, which may introduce an extra source of noise.

The results of the two time periods under study show a remarkably similar behaviour. Therefore, throughout this section we will mainly show the numerical values of the time period 1140–1200, for the sake of readability. In the supplementary material the reader can find a systematic analysis of the two periods under study.

The backbone exhibits high levels of clustering

The statistical analysis of networks shows remarkable degree of clustering at all levels of description. In the period 1140–1200, the average degrees for the various layers of the MPN are $\langle k \rangle_F = 18.15$, $\langle k \rangle_C = 16.15$, and $\langle k \rangle_T = 33.12$ and the clustering coefficients are remarkably high if we take into account these connectivities: $C_F = 0.235(0.037)$, $C_C = 0.235(0.06)$, and $C_T = 0.354(0.04)$. Numbers in brackets correspond to the expected value of the clustering coefficient in an ensemble of random networks having the same size and degree distribution than the real ones, see methods and S1 File. The intersection networks show a slight decrease on the number of nodes (see Table S1,S2 in S1 File) and smaller average degrees: $\langle k \rangle_{FC} = 6.27$, $\langle k \rangle_{FT} = 5.21$, $\langle k \rangle_{TC} = 7.05$, and most pronounced, $\langle k \rangle_{FCT} = 3.89$, as expected. Although the average degree is lower than in the MPNs, the clustering coefficients still show remarkably high values, especially when compared with the randomized values, $C_{FC} = 0.198(0.020)$, $C_{FT} = 0.249(0.009)$, $C_{TC} = 0.297(0.017)$, and $C_{FCT} = 0.197(0.006)$. The persistence of the clustering coefficient, even for $\mathcal{G}_{FCT}$, where the expected C for the randomized case almost vanishes, indicates that the mechanism of *triadic closure* [30–33] plays an important role in the dynamical formation of the backbone structure in social systems.

The G_K-sequence

We compute the G_K-decomposition sequence (see S1 File for details) and observe the following trends. We generally observe long G_K- decomposition sequences. The length of the decomposition sequence is the largest value of K for which G_K-core is not empty. For the different networks $\mathcal{G}_{FCT}, \mathcal{G}_{FC}, \mathcal{G}_{FT}, \mathcal{G}_{CT}, \mathcal{G}_F, \mathcal{G}_C$ and $\mathcal{G}_T$, these limit values are found at $K = 27, 38, 32, 42, 88, 111$ and, again 111, respectively.

In Fig. 2a the size of the *giant connected component* (GCC) [34] along the G_K-decomposition sequence is shown for the $\mathcal{G}_F$ network (black) -In a little abuse of notation, we refer to the GCC as the set of nodes that from a connected component significantly larger than the others, if there exist any. In our case, the G_K-cores generally show a single connected component. We observe that the G_K-decomposition sequence is longer than the one expected by chance, see Fig. 2a, (red). The situation for the traditional K-core is different, with a behaviour similar to the one expected by chance in all studied subgraphs, see Fig. 2d. Further, the evolution of the size GCC of the G_K-cores shows plateaus followed by
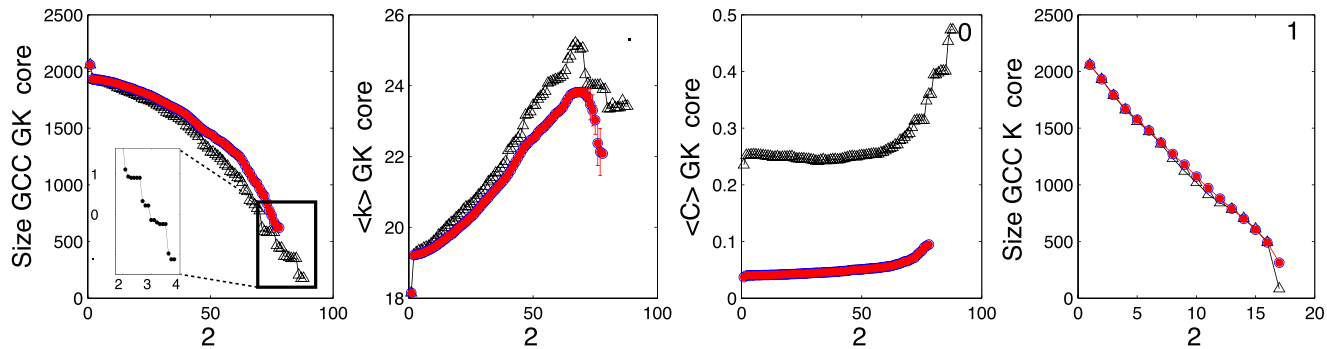
Fig. 2. Evolution of the topological indicators along the G_K-decomposition sequence for the $\mathcal{G}_F$ level of the MPN of the period 1140–1200. In a) we have the evolution of the size of the GCC of the G_K-core of the net (black) and its randomized counterpart (red). In the box inside the figure we highlight the evolution of the size of the GCC of the G_K-core at high K-levels, where flats regions followed by sudden decreases are observed. b) Evolution of the average degree of the G_K-core (black) and its randomized counterpart (red). c) Evolution of the average clustering coefficient of the net (black) against its randomized counterpart (red). Finally, in d) We plot the evolution of the GCC of the K-core of the net in terms of K (black) against its randomized counterpart (red). Observe that, for this latter plot, there are no significant statistical differences on the behaviour of the real graph when compared to the randomized one. The results for the random counterpart of the net have been obtained from an ensemble of 25 randomized versions of $\mathcal{G}_F$, see text and methods section.

doi:10.1371/journal.pone.0112606.g002

abrupt changes, which may depict different levels of core organization. On closer inspection, we find that often these changes signal the collapse of a cluster, which forms a cohesive community at certain level *K*, and which is completely absent at level $K+1$. The structure of the G_K-core just before a collapse represents one organizational level which is replaced by a deeper one, maybe with different topological and social characteristics. We observe that the length of the decomposition sequence strongly depends on the size of the network, a feature probably due to the power law degree distribution they exhibit. As shown in [21] for generic sequences of nested subgraphs, the depth of the decomposition sequence diverges for this kind of networks.

The evolution of the average degree $\langle k \rangle$ along the decomposition sequence for the $\mathcal{G}_F$ network is seen in Fig. 2b (black). We find significant differences between the social networks and their randomized counterparts (red). In most cases one observes that the average degrees along the decomposition sequence first increase with *K*, revealing a phenomenon which resembles the so-called *rich club* [17]. Here, elements of the G_K-core tend to be more connected among themselves than would be expected by chance. We find an exception in the $\mathcal{G}_T$ network where there are no significant differences between the real average degrees and those obtained after randomization. This increasing trend usually peaks and stops at deep levels, followed by a slight decrease at the deepest levels, see Fig. 2b. The increase is absent in standard models of random graph like the Erdös Rény [34] and Barabási-Albert [35] networks, see Fig. S1 of the S1 File. This means that the particular structure of the social network determines the functional form of this curve. Since the randomized ensembles also show an increasing trend of connectivity through the sequences, see Fig. 2b (red), one might expect that the degree distribution is partially responsible of the observed increase. Furthermore, the presence of high clustering could also be responsible for an additional increase

of the connectivity of the cores, thus explaining the deviation from their randomized counterparts.

Finally, the evolution of the clustering coefficient displays two clearly differentiated regions: At low and medium stages of the decomposition sequence it shows a more or less constant behaviour, followed by an increase at later stages of the sequence. This latter increase may also be the footprint of a rich-club phenomenon in the networks under study. It is worth to observe that along the decomposition sequence, the real values of the clustering coefficient are at least one order of magnitude higher than the expected by chance. In Fig. 2c we display the evolution of the clustering coefficient along the decomposition sequence for the $\mathcal{G}_F$ network.

Identification of characteristic K-levels and core communities through the M-core

In the previous section we pointed out that the evolution of the size of the G_K-core throughout the decomposition sequence eventually displays sudden decreases, and that such sharp decays might be related to massive collapses of communities the core. Such change might reveal different levels of core organization. How to identify such crucial levels and, therefore, communities inside the G_K-core? We assume that the cohesiveness of such communities leads to a high degree of transitivity between them, i.e., that the clustering coefficient inside such communities is exceptionally high. This intuition is supported by the extremely high clustering coefficient values found in the system under study, as we reported above. Moreover, we assume that the degree of transitivity between communities is very low namely, that connections between members of different communities are performed by simple links or by means of connector nodes. Under such defining assumptions of core community, the recently introduced M-core [36] plays a crucial role. The M-core is the *maximally induced subgraph in which each link participates at least in M triangles*. Therefore, the application of the M-core with $M=1$, $M=2$ over the G_K-cores will remove those links (and maybe some nodes) which do no participate in a highly clustered structure, eventually acting as bridges between communities. The *unconnected components that may emerge from the application of the M-core (M=1,2) to the G_K-core will be the core communities of our graph at level K*, see Fig. 3a,b, methods section and S1 File for a detailed information. For the sake of readability, let us refer to the M-core of the G_K-core as $M(G_K)$. As long as K increases, the number of components of $M(G_K)$ $(M=1,2)$ may fluctuate, thereby identifying different organizational levels within the core of the network. Such fluctuations, if any, will define different levels of core organization. In general, the deepest cores of the networks under study display only a single component, and we will put our focus on the last K by which $M(G_K)$ $(M=1,2)$ contains more than a single component. We will refer to this level of organization as the *characteristic K-level of organization*. It may happen that such a level does not exist, then we will conclude that for this network and under our assumptions, the G_K-core does not change dramatically its structure
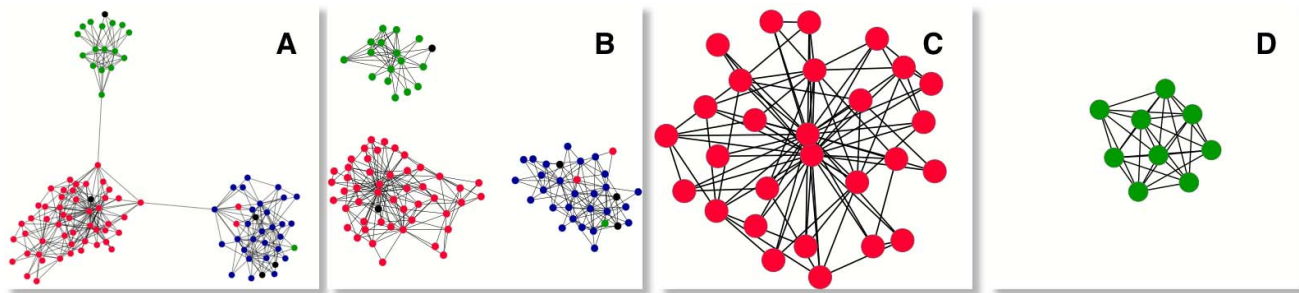
Fig. 3. National elites define topological communities at deep levels. The composition of the G_K-core in terms of nations reveals that the multiplex system is organised around the elites of the three existing nations, whose members are depicted with different colours (see text for the use of colours). We have a) the *characteristic* G_K for $\mathcal{G}_{FCT}$, where we find that the *K*-critical level is located at $K = 13$ b) after the application of the *M*-core ($M = 1$), three components appear isolated, to be identified as the three communities composing the G_K-core. Such communities are almost uniformly populated by members of the same nation. In c) we have the deepest G_K-core, which contains members of only one nation. Interestingly, the composition of the deepest *K*-core of the $\mathcal{G}_{FCT}$, $K = 7$, d), is absolutely different from the composition of the deepest G_K-core of the same net, located at $K = 26$, showing interesting qualitative differences between these two approaches of core extraction. All pictures belong to the period 1140–1200.

throughout the values of *K*. The rationale behind the definition of this characteristic level is clear: we want to study the structure of the core before the last reorganization, for it may contain many topological and properties absent in the deepest one. As we shall see, this methodology is able to perfectly identify core communities in our system, see Fig. 3a,b. It is worth to emphasise that randomized versions of the nets under study always display a single component and no communities –and, thus, no characteristic *K*-levels– can be identified.

With the *characteristic* G_K-core and the *deepest* G_K-core, we have two snapshots of the core organization, presumably depicting different structural features. The former represents a core structure which vanishes at deeper levels, the latter shows how the elements at the deepest level of description are organised. For the networks corresponding to the period 1140-1200, $\mathcal{G}_{FCT}, \mathcal{G}_{FC}, \mathcal{G}_{FT}, \mathcal{G}_{CT}, \mathcal{G}_F$, we got the following characteristic *K*-levels: $K = 13, 37, 23, 38$ and 5 respectively. $\mathcal{G}_C$ and $\mathcal{G}_T$ did not show any characteristic level. The networks obtained out of the intersection of MPN levels display a clearer core community structure and thus relevant characteristic levels can be identified. In the case of $\mathcal{G}_F$, the characteristic level is found at a very low *K*, so its statistical relevance is lower than the characteristic *K*-levels reported for the intersection nets.

The G_K-core and the elites of the social system

We can now characterize the individuals populating the cores of the various networks with a series of quantitative social indicators in the 'Pardus' society. These measure status, competence, social leadership, relevance and success of various kinds. In particular we use the following indicators, and we indicate how they appear in Table 1: *Experience* ($\langle$ Exp$\rangle$, in the table. Numerical indicator accounting for the experience of the player), *Activity* ($\langle$Act$\rangle$ in the table. Number of actions performed by the player), *Age* ($\langle$Age$\rangle$ in the table. Age in units of days after the player joined the game), *Wealth*, ($\langle$Wealth$\rangle$ numerical indicator

Table 1. Social indicators of the isolated groups of nodes.

	$\langle Exp \rangle$	$\langle Act \rangle$	$\langle Age \rangle$	$\langle Wealth \rangle$	gComp	FracL	$\langle GlobL \rangle$	N
$\mathcal{G}_{FCT}$								
Char. G_K	7.72×10^5	5.69×10^6	1.02×10^3	9.84×10^7	0.885	0.195	10.7	87
Hubs	1.01×10^6	6.86×10^6	1.08×10^3	1.23×10^8	0.933	**0.4**	11.4	15
Deep. G_K	9.78×10^5	5.96×10^6	1.09×10^3	1.14×10^8	0.962	0.154	11.3	26
Hubs	5.69×10^5	**7.39×10^6**	**1.2×10^3**	**3.03×10^8**	1	**1**	12	2
Deep. K-Core	7.18×10^5	6.23×10^6	1.09×10^3	**1.4×10^8**	0.889	0.111	11	9
All Net	4.86×10^5	3.88×10^6	857	4.87×10^7	0.875	0.165	7.64	1303
$\mathcal{G}_{FC}$								
Char. G_K	8.47×10^5	5.72×10^6	1.04×10^3	7.69×10^7	0.884	0.207	9.41	121
Hubs	1.32×10^6	6.96×10^6	**1.15×10^3**	1.24×10^8	0.778	0.333	12.6	9
Deep. G_K	8.07×10^5	5.59×10^6	1.01×10^3	6.37×10^7	0.882	0.235	8.69	85
Hubs	1.53×10^6	6.84×10^6	1.13×10^3	7.26×10^7	0.714	0.143	12.7	7
Deep. K-Core	9.4×10^5	6.03×10^6	1.01×10^3	6.66×10^7	0.882	0.329	9.5	76
All Net	4.69×10^5	3.72×10^6	842	4.35×10^7	0.871	0.154	7.4	1600
$\mathcal{G}_{FT}$								
Char. G_K	8.48×10^5	5.77×10^6	1.05×10^3	8.94×10^7	0.892	0.169	10.6	83
Hubs	**1.34×10^6**	**7.37×10^6**	1.13×10^3	**1.8×10^8**	0.889	0.333	12.1	9
Deep. G_K	9.2×10^5	5.87×10^6	1.11×10^3	1.1×10^8	0.935	0.194	11.3	31
Hubs	5.69×10^5	**7.39×10^6**	**1.2×10^3**	**3.03×10^8**	1	**1**	12	2
Deep. K-Core	7.18×10^5	6.23×10^6	1.09×10^3	1.4×10^8	0.889	0.111	11	9
All Net	4.76×10^5	3.77×10^6	869	4.46×10^7	0.872	0.143	7.56	1660
$\mathcal{G}_F$								
Char. G_K	1.9×10^5	1.88×10^6	608	1.86×10^7	0.86	0.0457	6.08	328
Hubs	5.05×10^5	4×10^6	925	4.65×10^7	0.87	0.155	7.61	1585
Deep. G_K	7.57×10^5	5.34×10^6	1.05×10^3	5.96×10^7	0.877	0.175	7.33	171
Hubs	**1.39×10^6**	6.68×10^6	**1.15×10^3**	7.8×10^7	**0.6**	0	**12.8**	5
Deep. K-Core	1.0×10^6	6.12×10^6	1.08×10^3	6.83×10^7	0.88	0.253	9.11	83
All players	4.3×10^5	3.5×10^6	841	3.96×10^7	0.87	0.12	7.51	2059

We show the scores for the cores of the $\mathcal{G}_{FCT}$, $\mathcal{G}_{FC}$, $\mathcal{G}_{FT}$ and $\mathcal{G}_F$ networks. 'Char. G_K' refers to the connectors of the *Characteristic* G_K, 'Hubs' below it refers to *Hubs of the Characteristic* G_K. 'Deep. G_K' refers to the connectors of the *Deepest* G_K. 'Hubs' below it refers to *Hubs of the Deepest* G_K. Deep. K-core refers to the nodes of the *Deepest K-core*. 'All net' refers to all players belonging to the net whose results for the different cores is shown immediately above. We highlighted in boldface the two highest average score for each indicator.

accounting for the wealth of the player within the game), *Fraction of leaders* (FracL, in the table. Fraction of players who are leaders in some aspect in a given subgroup of the society at the local level), and *Global leadership* ($\langle$GlobL$\rangle$ in the table. Numerical indicator evaluating the degree of leadership of the player). For detailed information about the definition of these indicators, see S1 File. We finally checked the *gender composition*, the fraction of male/female players in the core. We classify the nodes in the core whether they are a hub or a connector, and

present results accordingly. We also computed the scores obtained by the members belonging to the deepest K-core, of each studied graph. In Table 1 we show the scores from four networks $\mathcal{G}_{FCT}, \mathcal{G}_{FC}, \mathcal{G}_{FT}$ and $\mathcal{G}_F$, see S1 File for Tables with all social indicators over core subgraphs obtained from all networks belonging to the two periods under study.

The combination of the filtering provided by the intersection plus the G_K-core extraction clearly identifies the structured groups of players having the highest indicators of social performance and influence. Although, as we pointed out above, there is no null model for an elite detection, one can analyse how relevant are the nodes of the topologically isolated graphs within the collection of raw values of performance indicators belonging to all players of our MPN. Indeed, let us rank all players of the MPN with respect to their performance in a given indicator and then take the 10% best performers of such indicator. Then, to check if the nodes of our subgraphs are among the best performers we compare the actual number of members which belong both to a given G_K-core and to this top-10% set of players against the expected number of players belonging to the G_K-core who also belong to this top-10% set. What we observe is that, both for wealth and global leadership, the actual number of players of a given G_K-core which belong to the set of top-10% best performers scales up to 5 times the expected one, which shows that there is a strong relation between good performance within the society and being member of the G_K-core. In Fig. 4 we show the ratio between the actual number of members of the G_K-core belonging to the top-10% against the expected value. We show the evolution of such ratio for the two periods under study for *global leadership*, Fig. 4a, *Wealth*, Fig. 4b, *Activity*, Fig. 4c, and *Experience*, Fig. 4d. All plots show an increasing trend which stops around the characteristic K-level. Beyond this, the trend flattens and becomes stable, due to the very tiny variations suffered by the G_K-core at these levels, until it completely collapses.

In table 1 we highlighted in Boldface the two highest average scores for the following sets of nodes: Connectors of the G_K-core at the characteristic K-level, Hubs of the G_K-core at the characteristic level, Connectors of the deepest G_K-core, Hubs of the deepest G_K-core and the scores of the players of the whole network. We show the results for $\mathcal{G}_{FCT}, \mathcal{G}_{FC}, \mathcal{G}_{FT}$ and $\mathcal{G}_F$ for the period 1140-1200. In tables S1 and S2 of the S1 File the reader will find an exhaustive analysis of all the nets belonging to the two periods under study. Interestingly, the highest scores of a given network are not necessarily found at the deepest level of the decomposition sequence, but are usually found in the identified characteristic K-level, as seen in Table 1 in *Experience* in $\mathcal{G}_{FCT}$ and *Wealth* in $\mathcal{G}_{FC}$. This happens even though the number of players belonging to the characteristic K-level is substantially larger than the number of players populating the deepest G_K-core.

We finally check if the membership to the connector set of a G_K-core implies a distinction with respect to those players whose connectivity patterns are comparable. Specifically, we refer to individuals having the same degree than a given connector but not being members to the connector set of G_K. Suppose that an individual v_i is a connector in the characteristic K-level of $\mathcal{G}_{FCT}$, ($K=13$, for
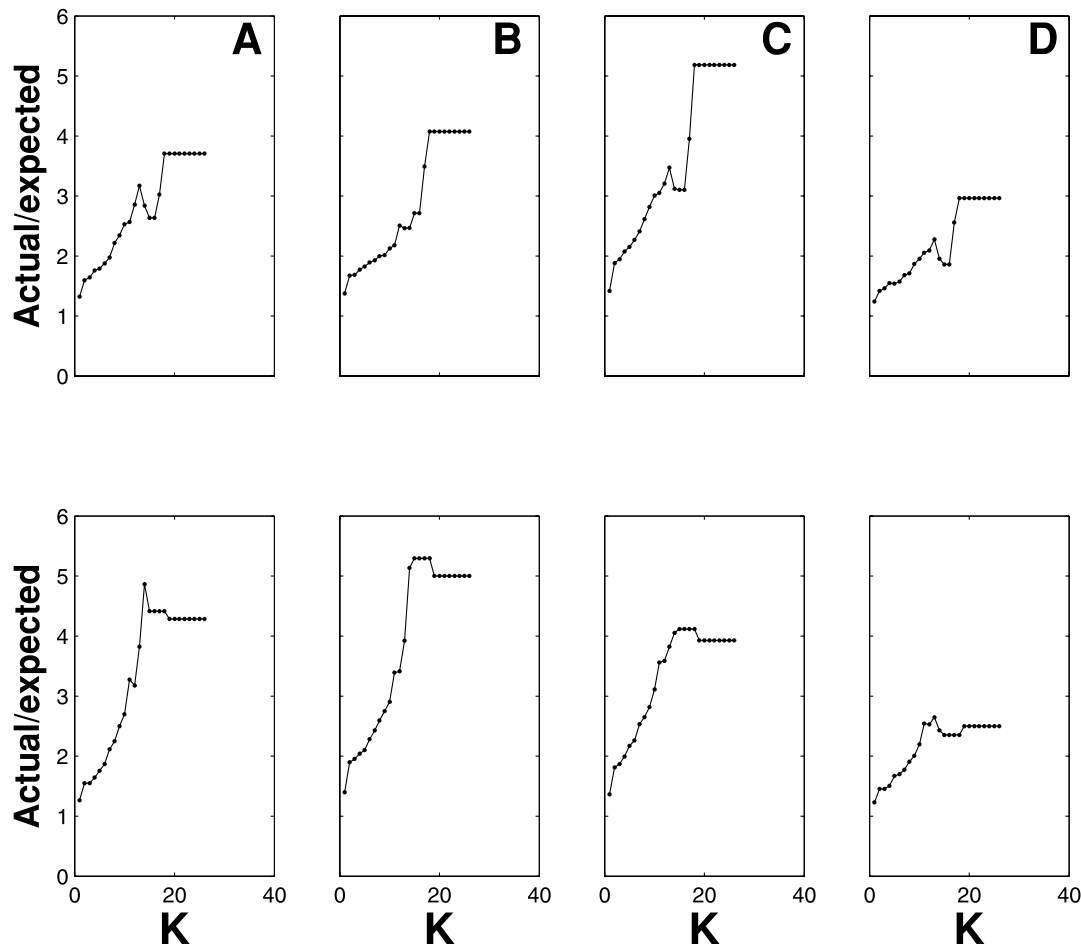
Fig. 4. Overabundance of members of the G_K-core in the set of the top-10% best performers of the game. In these plots we show the evolution along the G_K-decomposition sequence of the quotient between the actual number of members belonging to the G_K-core which also belong to the set of the top-10% best performers of a given indicator against the expected number of them in case they are spread randomly. On top we have the results for the period 756–856 and at the bottom we have the results for the period 1140–1200, both for the $\mathcal{G}_{FCT}$ networks of their respective periods. We plot this ratio for a) Wealth, b) Global leadership, c) Activity and d) Experience. All of them show an overabundance of members of the G_K-core, showing an intrinsic relation between better social performance and deep G_K-core membership. It is worth to observe i) the clear overabundance of members of the G_K-core within the set of the top 10% in any indicator and ii) the change of the trend after the characteristic K-level, which is $K = 16$ for the $\mathcal{G}_{FCT}$ of the period 796–856 and $K = 13$ for the period 1140–1200.

the period 1140-1200) with a degree in the $\mathcal{G}_{FCT}$ network of k_i. Now take all individuals in $\mathcal{G}_{FCT}$ whose degree is equal to k_i but who *do not* belong to the characteristic G_K of this net. We observe that the relative performance of connectors with respect to those associated non-connectors of same degree is about $20-40\%$ higher, in particular: $\langle Exp.\rangle_{G_K}/\langle Exp.\rangle_{\text{not}-G_K} \approx 1.42$, $\langle Act.\rangle_{G_K}/\langle Act.\rangle_{\text{not}-G_K} \approx 1.3$, $\langle Age\rangle_{G_K}/\langle Age\rangle_{\text{not}-G_K} \approx 1.2$ and $\langle Wealth\rangle_{G_K}/\langle Wealth\rangle_{\text{not}-G_K} \approx 1.3$. These results point to the fact that to belong to the G_K-core structure increases the chances of having high scores of social performance. In some cases, we observe that the performance of connectors of the deepest G_K-core is still higher than the one exhibited by the members of the

K-core, see, for example, $\langle Exp.\rangle$ for $\mathcal{G}_{FCT}$ in Table 1 and S1 File. Therefore, connectors, although in general they perform worse than hubs in the G_K-cores, could constitute a secondary elite, which presumably takes advantage of the knowledge of the underlying net of relations defining the dynamics of the social system.

G_K-core clusters identify national elites/sharp reorganization at deep levels

We finally look at the national composition of the cores. Players usually belong to one of three 'factions' existing in the game, which are the equivalent of countries or nations. These nations are labeled as 'nation 1', 'nation 2' and 'nation 3', associated to colours red, green and blue, respectively, in Figs. 3 and 5. Players shown in black are not associated to any nation. Over all the population of the *Artemis* universe, the fraction of players in each nation is 0.34, 0.27 and 0.21, for nations $1-3$, respectively. Players not associated to any nation represent a fraction of 0.13 of all players.

Along the G_K-decomposition sequence of all studied networks, the nation composition of the G_K-cores displays two well differentiated regions. At lower levels of K, the national composition of the G_K-core is close to the one corresponding to the whole society. At high K-levels, G_K-cores are populated only by members of a single nation. The shift between these two qualitatively different core organizations is abrupt, and occurs right after the characteristic K-level. This behavior can be clearly seen in Fig. 5a,c, where we plot the evolution of the national composition of G_K-cores along the G_K-decomposition sequence of $\mathcal{G}_{FCT}$ belonging to the two periods under study. The evolution of the national composition of the K-core also show a similar behaviour, although less abrupt and only at the very late stages of the K-core-decomposition sequence, see Fig. 5b,d.

The application of the M-core ($M=1$) over the G_K-core shows that the elites of the three nations are clearly identified as clusters at the characteristic K-level. This can be seen in Fig. 3a,b, where we have the G_K-core $\mathcal{G}_{FCT}$ at the characteristic K-level and the $M(G_K)$. As we can see, the proposed method combining the G_K-core and the M-core perfectly identifies three communities belonging to the three existing nations. Interestingly, the cohesion of the entire core structure across nations is assured only by connectors. At deeper K-levels, only members of one nation populate the G_K-core, forming a compact cluster with no community differentiation, see Fig. 3c. The deepest K-level of the K-core is also populated by individuals belonging all of them to the same nation, see Fig. 3d. It is worth to remark that, against intuition, the national cluster isolated by the deepest K-core differs completely from the one isolated by the deepest G_K-core. Finally, it is worth to mention that 10 of the 13 identified hubs of the characteristic G_K-core of $\mathcal{G}_{FCT}$ have a specific leadership role, whereas only 1 of the 9 members of the deepest K-core does.
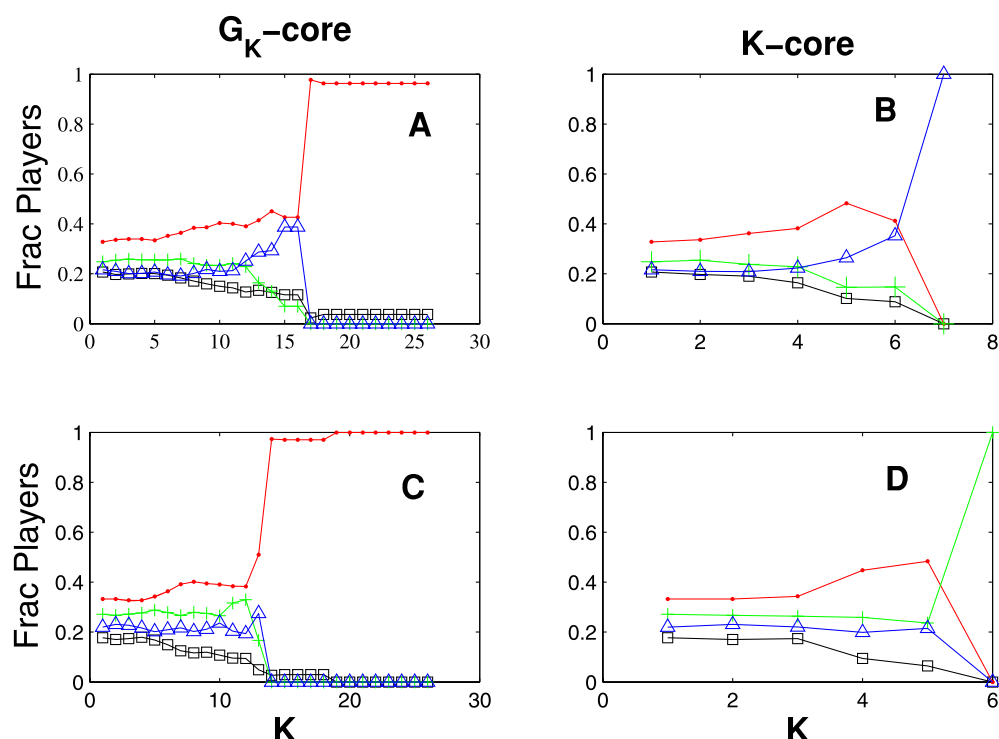
Fig. 5. Sharp transitions at the core organization of social networks. The value at $K = 1$ belongs to the composition of the society at the time period under study. On top a) we have the nation composition of the G_K-core and c) the K-core as a function of K for the $\mathcal{G}_{FCT}$ network corresponding to the period 796–856. At the bottom b) we have the nation composition of the G_K-core and (d) the K-core in terms of K for the G_{FCT} network corresponding to the period 1140-1200. Colours depict the different nations. As long as K increases, the composition of the cores in terms of nationalities is more or less stationary, with values close to the ones we find in the whole system. At certain K -right after the characteristic K- an abrupt change is observed a) for the and b), and the composition of the cores becomes uniformly populated by only one nation. The same phenomenon is observed when looking at the K-core decomposition sequence, although less pronounced. Notice that the deep G_K-cores isolated the same nation cluster in both periods (the 'red' nation), whereas the K-cores didn't.

Discussion

The aim of this study was to propose a topological method to detect the elites in a social system. We define elites not only as the set of highly connected individuals within a society, but as the set of highly connected ones *together* with their connectors in a network whose links depict multiple relations, like personal, communication or trade ones. Those elites are, presumably, strategically located at the core of the multiplex system defined by the society. To identify the elite cores, we suggest an algorithm that is similar in spirit to the traditional *K*-core, but that leads to entirely different compositions of the resulting core, which we called the *generalised K*-core. As a test system we used the human society of players of the MMOG Pardus, which not only provides the networks of various social interactions [7, 23–27], but also contains quantitative information of how individual players perform socially within the society in terms of leadership, wealth, social status among other skills, in which elite members are expected to

score exceptionally high. We find that elite structures are formed by hubs connected either directly or through connectors, generally at deep levels of the core (large K). Hubs of these core subsystems display the highest scores on social relevance, and this is especially true for the backbone network and for the networks obtained out of the intersection of two levels of the MPN, specifically, of friendship and communication levels, and of friendship and trade levels. In addition, we could show that connectors within the G_K-core perform consistently worse than hubs, however, we collected evidence pointing to the fact that connectors clearly socially outperform individuals (matched for their degree) that are not part of the G_K-core. This indicates that connectors could constitute something like a 'secondary' elite within the system, taking advantage of the knowledge they have of the underlying network of social relationships. In terms of national composition and core community structure, we have seen that a combined strategy including the use of the recently introduced M-core and the G_K-core clearly detects the clusters belonging to the elites of the three nations present in the game, thereby providing a new tool for community detection focused on the core properties of the net. Reorganization of the national composition of the cores happens in sharp bursts, rapid changes which are the footprint of the collapse of clusters within the core from one level K to another. In all performed analysis, it is worth mentioning the low performance of the K-core, when compared to the G_K-core to identify those leading subsets of individuals. We finally point out that, in spite of their low average degree, in all of the studied networks we found a remarkable level of clustering, which we attribute to the process of triadic-closure that seems to be a major driving force in the dynamics of social network formation [7, 30–33].

The presented results suggest that the subgraphs isolated by means of the G_K-core actually correspond to the way elites interact and define cohesive subgroups. In more general terms, further works could explore the role of connector nodes in terms of information flow within networks or their presumably relevant role when a dynamical process is defined over the network. It is reasonable to think that the combination of both low connectivity and their role of hinge between clusters may provide them a predominant role in terms of dynamic organization within the network. The proposed method could lead to a wide range of more general applications, such as network visualization or as a community detection algorithm.

Materials and Methods

Randomisation of Networks

Random ensembles of a given network $\mathcal{G}$ have been obtained after a rewiring process which keeps the degree of each node invariant. For a real network $\mathcal{G}$, we created 25 randomized versions by applying the rewiring operation 100 times the number of links of $\mathcal{G}$.

Intersection of different levels of the multiplex system

We formally refer to multiplex networks (MPNs) as $\mathcal{M}$, and to single graphs as $\mathcal{G}$. In a multiplex graph, $\mathcal{M}$, the set of nodes $V = \{v_1,...,v_n\}$ can be connected by different types of relations or links $\mathbf{E} = \{E_{\alpha_1},...,E_{\alpha_M}\}$, $E_{\alpha_k} = \{e_i(\alpha_k),...,e_m(\alpha_k)\}$. The whole multiplex is thus described by

$$\mathcal{M} = \mathcal{M}(V,\ E_{\alpha 1} \times ... \times E_{\alpha M}).$$

Let $E' = \{E_{\alpha_i},...,E_{\alpha_k}\}$, $E' \subset \mathbf{E}$, be a subset of the overall type of potential relations that can exist between two nodes, thereby redefining the concept of *link* as a collection of relations that relate two given nodes, instead of a single type of relation. We define the E'-intersection network, $\mathcal{G}_{E'}$ as

$$\mathcal{G}_{E'} = \mathcal{G}\left(V,\ \bigcap_{E_{\alpha_i} \in E'} E_{\alpha_i}\right).$$

In this network, links connect those pairs of nodes which are connected through, at least, links of type $E_{\alpha_i},...,E_{\alpha_k}$.

The generalised *K*-core

The *generalised K-core subgraph*, $G_K(\mathcal{G})$ of a given graph $\mathcal{G}$ is the maximal induced subgraph in which every node is either a hub with a degree equal or higher than K, or a connector that – regardless of its degree – connects at least 2 hubs with degree equal or higher than K. It can be obtained through a recursive pruning process. Starting with graph $\mathcal{G}$ we remove all nodes $v_i \in \mathcal{G}$ satisfying that: (1) its degree is lower than K *and* (2) at most one of its nearest neighbors has a degree equal or higher than K. We iteratively apply this operation over a finite graph $\mathcal{G}$ until no nodes can be pruned, either because the G_K-core is empty or because all nodes which survived the iterative pruning mechanism cannot be removed following the above instructions. The graph obtained after this process is the *generalised K-core subgraph*. Note that, for any finite graph, there exists a K^* by which even though $G_{K^*} \neq \varnothing$, $(\forall K > K^*)\ G_K(\mathcal{G}) = \varnothing$. We refer to $G_{K^*}(\mathcal{G})$ as the *deepest G_K-core* of the network $\mathcal{G}$, see S1 File for the algorithm.

The standard K-core is obtained by means of an iterative algorithm like the one shown above. The step of the algorithm consists in removing nodes whose degree is lower than K. This is performed iteratively until there are no more nodes to prune, see S1 File.

Finally, the M-core is obtained by means of an iterative algorithm like the ones shown above. The step of the algorithm consists in removing *links* participating in less than M triangles. Again, this is performed iteratively until there are no more nodes to prune, see S1 File.

Identifying levels of organization at the core

The definition of level of organization is based on the presence of highly clustered communities in the G_K-core and its eventual collapse when K increases. Specifically, given a graph $\mathcal{G}$:

- Compute its G_K-core
- Compute its M-core with $M=1$ over the G_K-core and check if the subgraph contains more than a single component. If not, compute the M-core ($M=2$) over the G_K-core and check if it contains more than a single component.
- Components of the $M(G_K)$ are the core communities at level K of the G_K-core.
- If the $M(G_K)$ with $M=1,2$ contains a different number of components than $M(G_{K+1})$ ($M=1,2$), K is a characteristic level of organization.

Throughout the paper we have been focused on the characteristic level of organization defined by the *largest K by which* $M(G_K)$, ($M=1,2$) *contains more than single component*. At deep levels, all the studied $M(G_K)$'s contain only a single component. Furthermore, it may happen that G_K itself contains more than a single component. This does not change the algorithm for characteristic K-level identification.

Supporting Information

S1 File. Rigorous definition of the algorithms. Study of the behaviour of standard models of networks. Systematic analysis of the topological properties studied in the main text through all networks under study. Table S1 and Table S2 of average social indicators for all the studied subgraphs for both periods under study.
doi:10.1371/journal.pone.0112606.s001 (PDF)

S1 Data. Data used to generate the results of the paper. Includes: A table of indicators of social performance for each player in each period under study and the three networks of the multiplex system for each period under study.
doi:10.1371/journal.pone.0112606.s002 (ZIP)

Acknowledgments

Author Contributions

Conceived and designed the experiments: BC-M ST. Analyzed the data: BC-M BF. Contributed reagents/materials/analysis tools: BC-M. Wrote the paper: BC-M BF ST.

References

1. **Mills CW** (1956) *The power elite* (Oxford University Press, Oxford).

2. **Mills CW** (1958) The structure of power in american society. The British Journal of Sociology 9(1):29–41.

3. **Keller S** (1963) *Beyond the Ruling Class. Strategic Elites in Modern Society* (Random House, New York).

4. **William FG** (1967) *Who rules America*? (McGraw-Hill).

5. **Bottomore T** (1993) *Elites and society, second ed* (Routledge, London).

6. **Mucha PJ, Richardson T, Macon K, Porter MA, Onnela JP** (2010) Community structure in time-dependent, multiscale, and multiplex networks. Science 328: 876–878.

7. **Szell M, Thurner S** (2010) Measuring social dynamics in a massive multiplayer online game. Social Networks 39: 313–329.

8. **Nicosia V, Bianconi G, Latora V, Barthelemy M** (2013) Growing multiplex networks. Phys Rev Lett 111: 058701.

9. **Wasserman S, Faust K** (1994) *Social Network Analysis* (Cambridge University Press).

10. **Freeman LC** (1978) Centrality in social networks conceptual clarification. Soc Networks *p 215*.

11. **Friedkin NE** (1984) Structural Cohesion and Equivalence Explanations of Social Homogeneity. Sociological Methods & Research 12: 235–261.

12. **Seidman SB** (1983) Network structure and minimum degree. Social Networks 5: 269–287.

13. **Bollobás B** (1984) The evolution of sparse graphs. Graph Theory and Combinatorics, Proc Cambridge Combinatorial Conf in honor to Paul Erdös, Academic press pp 35–57.

14. **Dorogovtsev SN, Goltsev AV, Mendes JFF** (2006) k-core organization of complex networks. Phys Rev Lett 96: 040601.

15. **Harary F, Ross IC** (1957) A procedure for clique detection using the group matrix. Sociometry 20: 205–215.

16. **Bianconi G, Marsili M** (2006) Emergence of large cliques in random scale-free networks. EPL (Europhys Lett) 74: 740.

17. **Colizza V, Flammini A, Serrano MA, Vespignani A** (2006) Detecting rich-club ordering in complex networks. Nat Phys 2: 110–115.

18. **Girvan M, Newman MEJ** (2002) Community structure in social and biological networks. Proc of the Natl Acad of Sci 99: 7821–7826.

19. **Palla G, Derenyi I, Farkas I, Vicsek T** (2005) Uncovering the overlapping community structure of complex networks in nature and society. Nature 435: 814–818.

20. **Corominas-Murtra B, Valverde S, Rodríguez-Caso C, Solé RV** (2007) K-scaffold subgraphs of complex networks. EPL (Europhys Lett) 77: 18004.

21. **Corominas-Murtra B, Mendes JFF, Solé RV** (2008) Nested subgraphs of complex networks. J of Phys A: Math and Theor 41: 385003.

22. **Castronova E** (2005) *Synthetic Worlds: The Business and Culture of Online Games* (Univ of Chicago Press, Chicago).

23. **Szell M, Lambiotte R, Thurner S** (2010) Multirelational organization of large-scale social networks in an online world. Proc of the Natl Acad of Sci 107: 13636–13641.

24. **Szell M, Thurner S** (2012) Social dynamics in a large-scale online game. Adv in Complex Sys 15: 1250064.

25. **Szell M, Sinatra R, Petri G, Thurner S, Latora V** (2012) Understanding mobility in a social petri dish. Sci Rep 2: 457.

26. **Thurner S, Szell M, Sinatra R** (2012) Emergence of good conduct, scaling and zipf laws in human behavioral sequences in an online world. PLoS ONE 7:e29796.

27. **Szell M, Thurner S** (2013) How women organise social networks different from men: gender-specific behavior in large-scale social networks. Sci Rep 3: 1214.

28. **Fuchs B, Thurner S** (2014) Behavioral and Network Origins of Wealth Inequality: Insights from a Virtual World. PLoS ONE 9(8): e103503. doi:10.1371/journal.pone.0103503

29. **Bianconi G** (2013) Statistical mechanics of multiplex networks: Entropy and overlap. Phys Rev E 87: 062806.

30. **Rapoport A** (1953) Spread of information through a population with socio-structural bias: I. Assumption of transitivity. Bull of Math Biol 15: 523–533.

31. **Granovetter M** (1973) The Strength of Weak Ties. Am J Sociol 78: 1360–1380.

32. **Davidsen J, Ebel H, Bornholdt S** (2002) Emergence of a small world from local interactions: Modeling acquaintance networks. Phys Rev Lett 88: 128701.

33. **Klimek P, Thurner S** (2013) Triadic closure dynamics drives scaling laws in social multiplex networks. New J of Phys 15: 063008.

34. **Newman ME, Strogatz SH, Watts DJ** (2001) Random graphs with arbitrary degree distributions and their applications. Phys Rev E *64*.

35. **Barabási AL, Albert R** (1999) Emergence of scaling in random networks. Science 286: 509–512.

36. **Colomer-de-Simón P, Serrano MÁ, Beiró MG, Alvarez-Hamelin JI**, **Boguñá** (2013) Deciphering the global organization of clustering in real complex networks. Scientific Reports 3: 2517 doi:10.1038/srep02517