Detection of the elite structure in a virtual multiplex social system by means of a generalized $K$-core

Bernat Corominas-Murtra$^1$, Benedikt Fuchs$^1$ and Stefan Thurner$^{1,2,3}$

$^1$ Section for Science of Complex Systems; Medical University of Vienna, Spitalgasse 23; A-1090, Austria
$^2$Santa Fe Institute, 1399 Hyde Park Road, New Mexico 87501, USA
$^3$ IIASA, Schlossplatz 1, A-2361 Laxenburg; Austria

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 society comprised of 420,000 people engaged in 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 generalized $K$-core identify those individuals that are high social performers in terms of a series of indicators that are available in the game. In addition, the elites of the different ‘nations’ present in the game are perfectly identified as modules of the generalized $K$-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.

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. 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). 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 in the various networks of the MPN, which enables them to 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. In this view elites are ‘core-communities’ that, to a certain extent, organize the whole topology of social interactions in a social system. 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. 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, the clique identification or the rich club analysis, among other general methods of cohesive subgroup identification. 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 generalized $K$-core is obtained by an iterative method inspired both by the so-called $K$-scaffold and the $K$-core. Specifically, the generalized $K$-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 tra-
A multiplex network (MPN) is a network composed of several layers of different relations among nodes. One can define link-sets that capture specific social interactions, such as communication, friendship, trading, enmity, attack, and revenge. These networks are tightly related and mutually influence each other as they have been systematically explored and quantified in [1, 23–27]. 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 \[ M(t) = M(V, E_F \times E_C \times E_T) \]. We restrict ourselves to a subset of the whole player base where the three networks \( G_F, G_C \) and \( G_T \) share the same set of 2059 nodes, \( V \). For simplicity we carry out our analysis at a particular point in time, \( t = 1200 \) in units of days since beginning of the game in Sept 2004. We drop the time index.

### B. Core extraction and elite identification

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 \[ G_{FC} = G_F \cap G_C, \ G_{FT} = G_F \cap G_T, \ G_{CT} = G_C \cap G_T \] and \[ G_{FCT} = G_F \cap G_C \cap G_T \]. The links of these networks, often called multi-links [28], 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 \( G_{FCT} \), 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 interaction graphs.

The core organization of \( G \) 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,

\[ \ldots \subseteq G_K(G) \subseteq G_{K-1}(G) \subseteq \ldots \subseteq G_2(G) \subseteq G \].

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 extent it relates to properties that are expected from 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 organize; 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 countries 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.

I. RESULTS

We extract the mentioned seven networks from the Pardus data, in the same way as described in [17–24].

A. The backbone exhibits high levels of clustering

The statistical analysis of networks shows remarkable degree of clustering at all levels of description. The average degrees for the various MPNs 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 appendix. The intersection networks show a slight decrease on the number of nodes (see Table 1 in SI) 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 \( G_{FCT} \), where the expected \( C \) for the randomized case almost vanishes, indicates that the mechanism of triadic closure [29–32] plays an important role in the dynamical formation of the backbone structure in social systems.

B. \( G_K \)-sequences and identification of characteristic \( K \)-levels

We compute the \( G_K \)-decomposition sequence (see appendix for details) and observe the following trends. We generally observe long \( G_K \)-decomposition sequences. The length of the decomposition sequence is denoted by \( K^* \), which is the largest value of \( K \) for which \( G_K \) is not empty. Thus we naturally refer to the graph \( G_K \) as the deepest \( G_K \)-core of \( G \). For the different networks \( G_{FCT}, G_{FC}, G_{FT}, G_{CT}, G_F, G_C \) and \( G_T \), these limit values are found at \( K^* = 27, 38, 32, 42, 88, 111 \) and, again 111, respectively.

The evolution of the average degree \( \langle k \rangle \) along the decomposition sequence for the \( G_F \) network is seen in Fig. 2a (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 \( G_T \) 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, at about \( K = 70 \), in Fig. 2a. The increase is absent in standard models of random graph like the Erdös Rényi [33] and Barabási-Albert [34] networks, see appendix. 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. 2a, red curve, 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.

In Fig. 2b the size of the giant connected component (GCC) [33] along the \( G_K \)-decomposition sequence is shown for the \( G_F \) network (black). We observe that the \( G_K \)-decomposition sequence is longer than the one expected by chance, see Fig. 2b, (red). The situation for the traditional \( K \)-core is different, with a behavior similar to the one expected by chance in all studied subgraphs, see Fig. 2c. Further, the evolution of the size GCC of the \( G_K \)-cores shows plateaus followed by abrupt changes, which can be used to define characteristic levels of the network. 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 quantify the change of the size of the \( G_K \)-core along the decomposition sequence by computing the relative de-
crease of the size of its GCC from level $K$ to $K + 1$, which we call $\Delta f_K$, see methods and appendix. The size changes are shown in Fig. 2d, where big restructuring events appear as clear peaks. Among them, we study in more detail what happens at the level just before the second last sharp decay. We label this level by $\bar{K}$. With the $G_{\bar{K}}$-core and $G_{\bar{K}^*}$-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 organized. In Fig. 2b and 2d we highlight such an abrupt collapse of the $G_K$-core at $K = 80$, which means that $\bar{K} = 79$. For the seven networks $G_{F^{CT}}, G_{F^{FC}}, G_{F^{CT}}, G_{F^{C}}, G_{C}$ and $G_{T}$, we obtain $\bar{K} = 13, 33, 19, 31, 79, 103$ and $99$, respectively.

C. The hubs of the $G_K$-core form 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: Experience (numerical indicator accounting for the experience of the player), Activity (number of actions performed by the player), Age (in units of days after the player joined the game), Wealth, (numerical indicator accounting for the wealth of the player within the game), Fraction of leaders (fraction of players who are leaders in some aspect in a given subgroup of the society), and Global leadership (numerical indicator evaluating the degree of leadership of the player). For detailed information about the definition of these indicators, see appendix. 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 ($G_{K}^{\text{Hub}}$), or a connector ($G_{K}^{\text{Con}}$), and present results accordingly. We also computed the scores obtained by the members belonging to the deepest $K$-core, $K^* C$, of each studied graph. In Table 1 we show the most relevant scores from three networks $G_{F^{CT}}, G_{F^{C}}$ and $G_{F}$, see appendix for a Table with all indicators over core subgraphs obtained from all networks. We have chosen Experience (Exp.), Wealth (Wealth), the fraction of players in the subgraph having a leading role in its own alliance (FracLead) and Global leadership (GlobLead) as the most relevant indicators of social relevance and influence.

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. Interestingly, the highest scores of a given network are not necessarily found at the deepest level of the decomposition, $G_K^*$, but are usually found in $G_{\bar{K}}^*$, as seen in Table 1 in Experience in $G_{F^{CT}}$ and Wealth in $G_{F^{C}}$ and $G_{F}$. This happens in spite the number of players belonging to $G_{\bar{K}}^*$ is substantially larger than the number of players populating $G_K^*$, see Table 1 of the appendix. This is another indicator why the choice of $\bar{K}$ as a relevant level is reasonable. Looking at Table 1 and Table 1 of the SI, we see that the subsets of isolated nodes displaying the best scores on social performance are the following:

FIG. 2: Statistical properties of the successive cores for the friendship network $G_{F}$: (a) Average degrees $\langle k \rangle$ of the $G_{F}$ as a function of $K$ (black), compared to the one expected for its randomized ensemble (red). A clear increasing trend is observed, peaking at around $K \approx 70$. (b) Size of the GCC of the $G_K$-core (black) and size of the GCC of the $G_K$-cores for the randomized ensemble (red). We see the formation of plateaus indicating characteristic levels of organization, which are clearly absent in the $G_K$-decomposition sequences of the randomized ensemble. The last non-empty $G_K$-core is found at $K^* = 89$. (c) The size of the successive subgraphs of the $K$-core decomposition sequence (black) and its randomized counterpart (red). Clearly, $K^*$ (the deepest $K$-core) is found at $K^* = 18$, and no significant deviations are observed with respect the randomized system. (d) Size changes of the GCC across $K$ enable to identify collapses of communities and thereby characteristic levels on the organization of the nets. The social network (black) shows distinct peaks which correspond to the sharp decays of in (b). Peaks depict the collapse and removal of a sub-community and a resulting change on the internal organization. In the second peak starting from the end we identify the characteristic level $\bar{K}$, which here is $\bar{K} = 79$, shown in (b) and (d). Randomized versions don’t show peaks (red).
The best scores is significantly worse than the one of the set of players containing no local leaders, see Table 1 and Ta-

There is an interesting deviation in the case of the core. The Gender composition of the cores shows a rela-

tively higher trend to the presence of males at high levels of $K$. Note that a fraction of 0.87 of players are males.

There is an interesting deviation in the case of the $G_F$ network, whose hubs at $G_K$-core contain only a fraction of 0.6 males; moreover, this is the only isolated subset of players containing no local leaders, see Table 1 and Table 1 of the appendix.

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 of a given connector but not being members to the connector set of $G_K$. Suppose that an individual $v_i$ is a connector in $v_i \in G^c_{\text{Con}}(G_FCT)$, with a degree in the $G_{\text{FCT}}$ network of $k_i$. Now take all individuals in $G_{\text{FCT}}$ whose degree is equal to $k_i$ but do not belong to the $G_K$ 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: $<\text{Exp.}>_{G_K}/<\text{Exp.}>_{\text{not}-G_K} \approx 1.42, <\text{Act.}>_{G_K}/<\text{Act.}>_{\text{not}-G_K} \approx 1.3, <\text{Age}>_{G_K}/<\text{Age}>_{\text{not}-G_K} \approx 1.2$ and $<\text{Wealth}>_{G_K}/<\text{Wealth}>_{\text{not}-G_K} \approx 1.3$. 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 $G_K$-core is still higher than the one exhibited by the members of the $K$-core, see, for example, $<\text{Exp.}>_{G_FCT}$ in Table 1 and Table 1 of the appendix.

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.
D. $G_K$-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 colors blue, yellow and red, respectively, in Figs. 3 and 4. Players shown in grey are not associated to any nation. Overall 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 in general abrupt, and occurs around $K$. This behavior can be clearly seen in Fig. 3a,c, where we plot the evolution of the national composition of $G_K$-cores along the $G_K$-decomposition sequence of $G_{CT}$ and $G_{FCT}$, respectively. We observe that the reported shift is extremely abrupt in the intersection networks and less pronounced especially in $G_P$ and $G_C$. The evolution of the national composition of the $K$-core also show a similar behavior, although less abrupt and only at the very late stages of the $K$-$C$-decomposition sequence, see Fig. 3b,d.

The topological analysis shows that the elites of the three nations are clearly identified as clusters at the $K$-level. This can be seen in Fig. 4a,b, where we have the $G_K$-core of both $G_{CT}$ and $G_{FCT}$ at their respective $K$-levels. Interestingly, in the case of $G_{FCT}$ the cohesion of the entire core structure across nations is assured only by connectors. We observe that the $K$-core somewhat identifies nation clusters but with much less clarity than the $G_K$-core, see Fig. 4c. At deeper $K$-levels, only members of one nation populate the $G_K$-core, forming a compact cluster with no community differentiation, see Fig. 4d, top. The deepest $K$-level of the $K$-core is also populated by individuals belonging all of them to the same nation, see Fig. 4d, bottom. Interestingly, 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 $G_K$-core of $G_{FCT}$ have a specific leadership role, whereas only 1 of the 9 members of the deepest $K$-core do.

II. 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 generalized $K$-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 $K$), shortly before the core breaks at $K^*$. Hubs of these core subsystems display the highest scores on social relevance, and this is especially true for the backbone network and for the intersection network of friendship and communication. 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 connec-
tors 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 from the underlying network of social relationships. In terms of national composition, we have seen that the $G_K$-core clearly detects the clusters belonging to the elites of the three nations present on the game. 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 \cite{7,21,32}.

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. The proposed method could lead to a wide range of more general applications, such as its use as a community detection algorithm for networks in general.

III. METHODS

A. Randomization 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 100 randomized versions by applying the rewiring operation 100 times the number of links of $\mathcal{G}$.

B. 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 $E = \{E_1,...,E_M\}$, $E_{a_k} = \{e_{i}(\alpha_k),...,e_{m}(\alpha_k)\}$. The whole multiplex is thus described by

$$\mathcal{M} = \mathcal{M}(V, E_{a_1} \times \ldots \times E_{a_M})$$

Let $E' = \{E_{a_1},...,E_{a_M}\}$, $E' \subset E$, 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

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

In this network, links connect those pairs of nodes which are connected through, at least, links of type $E_{a_1},\ldots,E_{a_k}$.

C. The generalized $K$-core

The generalized $K$-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 Generalized $K$-core subgraph. Note that, for any finite graph, there exists a $K^*$ by which even though $G_{K^*} \neq \emptyset$, $(\forall K > K^*) G_{K^*}(\mathcal{G}) = \emptyset$. We refer to $G_{K^*}(\mathcal{G})$ as the deepest $G_K$-core of the network $\mathcal{G}$, see appendix for the algorithm.

The standard $K$-core is obtained by means of an iterative algorithm like the one shown above. The stop 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 appendix.

D. Identifying levels of organization at the core

The definition of level of organization is based on the study of the first differences on the size of the successive $G_K$-cores defining the $G_K$-decomposition sequence of the graph $\mathcal{G}$. Let $|V_{G_K}|$ be the size (number of nodes) of the $G_K(\mathcal{G})$. Given the graph $\mathcal{G}$

$$\Delta f_K(\mathcal{G}) = 1 - \frac{|V_{G(K+1)}|}{|V_{G(K)}|},$$

i.e., the relative size of the first differences on the size of $G_K$. Peaks of this function correspond to the collapse of large communities of nodes, which is used as a marker of a potentially important change on the organization of the $G_K$-core.

Acknowledgements This work was supported by the Austrian Science Fund FWF under KPP23378FW, the EU LASAGNE project, no. 318132 and the EU MULTIPLEX project, no. 318132. BC-M thanks Andreu Corominas-Murtra for suggesting discussions.
IV. EXTRACTING THE CORE

A. Intersection of different levels of the multiplex system

Let us have a graph \( G(V, E) \) where \( V = \{v_1, ..., v_n\} \) is the set of nodes and \( E = \{e_1, ..., e_m\} \) the set of links connecting this nodes. Given a node \( v_i \) its degree is the number of first neighbors, or nodes a given node is linked to, to be written as \( k(v_i) \). The probability that a randomly chosen node has degree \( k \) is \( p(k) \). The first moment of the degree distribution gives us the average degree \( \langle k \rangle = \sum_k kp(k) \). Social systems are better described by means of multiplex graphs, which can be thought of as different graphs sharing the same set of nodes. In a multiplex graph, \( M \), the set of nodes \( V = \{v_1, ..., v_n\} \) can be connected by different types of relations or links \( E = \{E_{i1}, ..., E_{im}\} \), \( E_{ik} = \{e_{i1}(\alpha_k), ..., e_{im}(\alpha_k)\} \). The whole multiplex system is thus described by:

\[
M = M(V, E_{a1} \times ... \times E_{am}).
\]  (1)

In these networks, concepts such as degree distribution or average degree are relative to the type of relations (links) we are interested in. Now let \( E' = \{E_{i1}, ..., E_{ak}\} \), \( E' \subset E \), be a subset of the overall type of potential relations that can exist between two nodes. We define the \( E' \)-intersection network, \( G_{E'} \) as follows:

\[
G_{E'} = G \left( V, \bigcap_{E_{ai} \in E'} E_{ai} \right). \tag{2}
\]

In this network, links connect those pairs of nodes which are connected, at least, by links of type \( E_{a1}, ..., E_{ak} \). Links in \( G_{E'} \) are called multilinks.

B. The backbone of the multiplex system

A special and particularly interesting case of equation (2) is the graph of the intersection of all types of relations, \( G_I \) (to be named \( G_{FCT} \), in the main text), which depicts the backbone of the multiplex system depicted by \( M \), namely:

\[
G_I = G \left( V, \bigcap_{1 \leq i \leq M} E_{ai} \right). \tag{3}
\]

We point out that we have to be careful when choosing the different sets of links \( E_{a1}, ..., E_{am} \), since antagonistic relationships (such as enmity and friendship) can lead to empty intersections. We thereby restrict the definition of the intersection graph when this is performed over compatible sets of links.

C. The Generalized K-core subgraph

The Generalized K-core subgraph of a given graph \( G \), \( G_K(G) \) or \( G_K \)-core of \( G \), is the maximal induced subgraph within which every node is either a hub (its degree is equal or higher than \( K \)) or a connector (its degree is lower than \( K \) but it connects at least 2 hubs). Increasing the threshold \( K \) we obtain the decomposition sequence of \( G \) in terms of \( G_K \), namely:

\[
... \subseteq G_K(G) \subseteq G_{K-1}(G) \subseteq ... \subseteq G_2(G) \subseteq G.
\]

We will refer to the above sequence as the \( G_K \)-decomposition sequence of \( G \). The \( G_K \)-core of a given graph \( G \) can be obtained through an iterative pruning process: Suppose an operation \( H_K(G) \) by which we prune all the nodes \( v_i \in G \) satisfying both that

- its degree is lower than \( K \) and
- at most 1 of its nearest neighbors has degree equal or higher than \( K \).

If we iteratively apply this operation over a finite graph \( G \),

\[
H^N_K(G) = H_K \circ ... \circ H_K(G),
\]

we will reach a value, \( n = N \), by which (\( \forall M > N \))

\[
H^N_K(G) = H_K^N(G).
\]

We can take it as a definition of the generalized \( K \)-core, by saying that:

\[
G_K(G) = H_K^N(G). \tag{4}
\]

The equivalence between this definition and the one provided above can be easily checked: Indeed, on one hand, the algorithm itself forbids the presence of a node which is neither a hub or a connector, because, thanks to its iterative nature, it only stops when all surviving nodes satisfy the conditions to belong to the \( G_K \). On the other hand, we observe that the set isolated by the iterative algorithm is maximal: If a node (or a set of nodes) satisfies the conditions imposed by the algorithm, it is not pruned. We observe that, in any finite graph, \( \exists K^* \) by which although \( G_{K^*} \neq \emptyset \), \( (\forall K > K^*) G_{K}(\emptyset) = \emptyset \). We will refer to \( G_{K^*}(\emptyset) \) as the deepest \( G_K \)-core of \( G \).

Due to the potential richness of connectivity patterns that are allowed inside the \( G_K \)-core, we can categorize its nodes according to their topological roles:

- \( G_{K^{Con}}(G) \) is the set of nodes of \( G_K(G) \) whose degree is lower than \( K \), the \( K \)-connectors,
- \( G_{K^{Hub}}(G) \) is the set of nodes of \( G_K(G) \) whose degree is equal or higher than \( K \), the \( K \)-hubs.

• The set of \( K \)-critical connectors. A \( K \)-critical connector is a \( K \)-connector whose removal implies the breaking of the \( G_K(G) \) in two or more parts. We can analogously define the set of \( K \)-critical hubs.
D. The $K$-core subgraph

We end this section by describing the $K$-core subgraph. For the $K$-core definition and the exploration of its interesting properties, we refer the interested reader to [12–14]. The $K$-core of a given graph $G$, $KC(G)$, is the maximal induced subgraph whose nodes have degree at least $K$. It can be obtained through the application of an algorithm qualitatively close to the one described above by iteratively removing nodes whose degree is lower than $K$. The sequence

$$ \ldots \subseteq KC(G) \subseteq (K-1)K(G) \subseteq \ldots \subseteq 2C(G) \subseteq G. $$

is the $KC$-decomposition sequence, and the largest $K$ by which $KC(G) \neq \emptyset$, the deepest $K$-core, will be referred to as $K^\ast C(G)$.

V. MODEL NETWORKS

We now explore the behavior of the $G_K$ decomposition of two standard models of random graphs, namely, the Erdős-Rényi (ER) graph [33] and the Barabási-Albert (BA) graph [34]. For every type of graph we create an ensemble of 100 networks each, with $\langle k \rangle = 12$ in both the BA and the ER ensemble. We compute the evolution of the Giant Connected Component of all the non-empty $G_K$-cores and $K$-cores of the corresponding decomposition sequences and we plot them as a function of the threshold defined by $K$, see Fig. 5. For the BA scale-free networks, we observe a long decomposition sequence, thereby obtaining a picture of the core topology of the net at many different levels, see Fig. 5a. The behavior of the $G_K$-decomposition sequence for the ER ensemble shows that the $G_K$-core is either the whole graph or empty which can be due to the almost uniform degree distribution of this kind of graph see Fig. 1d. The behavior of the two ensembles is qualitatively similar under the $KC$-decomposition, showing an all-to-nothing transition at values close to $\langle k \rangle / 2$, see Fig. 1b,e. The average degree of the successive $G_K$-core subgraphs shows a slightly descending trend, whereas it remains constant in the case of the ER ensemble, mainly because, if the $G_K$-core is non-empty, it contains almost the whole graph. The counterintuitive decay in $\langle k \rangle$ for the BA ensemble can be explained by the increasing relative abundance of connectors against hubs within $G_K$ as long as $K$ increases.

VI. INDICATORS OF PERFORMANCE

We explored the behavior of 7 quantitative indicators of social performance within the ‘Pardus’ game (www.pardus.at):

- **Experience** is a numerical indicator accounting for the experience of the player, related to battles in which the player has participated, or the number monsters he/she killed.
- **Activity** is a numerical indicator related to the number of actions performed by the player.
• Age is the number of days after the player joined the game,

• Wealth, numerical indicator accounting for the wealth of the player within the game. Wealth accounts for cash money, value of the equipment the player owns within the game.

• Fraction of leaders fraction of players who are leaders in some aspect in a given alliance. Alliance should not be confused with nations. Alliances are small, organized groups of players. In the studied universe, we identify around \( \sim 140 \) different alliances. Every alliance has its own local leaders.

• Global leadership is numerical indicator evaluating the degree of leadership of the player. It is increased by doing missions, which are mainly transporting goods or killing monsters. The higher the Global leadership, the more powerful items may be bought – and the more missions are required to reach the next level. In general we can say that the higher this indicator, the more powerful and influential is the player within the whole society defined by the game.

• Gender composition evaluates the fraction of males within a given group of players.

In the table we show the scores of all these indicators for all 7 studied graphs at their respective \( K \) and \( K^* \)-levels. We distinguish between connectors and hubs. We compute theses social indicators for the \( K^* \)-core as well.

[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] Mucha PJ, Richardson T, Macon K, Porter MA, Onnela JP (2010) Community structure in time-dependent, multiscale, and multiplex networks. *Science* 328:876–878.
[5] Szell M, Thurner S (2010) Measuring social dynamics in a massive multiplayer online game. *Social Networks* 32:313–329.
[6] Nicosia V, Bianconi G, Latora V, Barthelemy M (2013) Growing multiplex networks. *Phys Rev Lett* 111:058701.
[7] Wasserman S, Faust K (1994) *Social Network Analysis* (Cambridge University Press).
[8] Freeman LC (1978) Centrality in social networks conceptual clarification. *Soc Networks* p 215.
[9] Corominas-Murtra B, Valverde S, Rodríguez-Caso C, Solé RV (2007) K-scaffold subgraphs of complex networks. *EPL (Europhys Lett)* 77:18004.
[10] Castronova E (2005) *Synthetic Worlds: The Business and Culture of Online Games* (Univ of Chicago Press, Chicago).
[11] 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.
[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 Erdős, 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] Barabási A, Ross IC (1957) A procedure for clique detection using the group matrix. *Sociometry* 20:205–215.
[16] Colizza V, Flammini A, Serrano MA, Vespignani A (2006) Detecting rich-club ordering in complex networks. *Nat Phys* 2:110–115.
[17] Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc of the Natl Acad of Sci* 99:7821–7826.
graphs with arbitrary degree distributions and their applications. Phys Rev E 64.

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

[35] 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 existence of unconnected components after the core extraction turns out to be negligible.
| \(G_{FC}^{\text{Con}}\) | \(G_{K}^{\text{Con}}\) | \(G_{\text{Hubs}}^{\text{Con}}\) | \(K^*C\) | all Net | \(G_{FC}^{\text{K}}\) | \(G_{\text{Hubs}}^{\text{K}}\) | \(K^*C\) | all Net | \(G_{FC}^{\text{K}}\) | \(G_{\text{Hubs}}^{\text{K}}\) | \(K^*C\) | all Net | \(G_{FC}\) | \(G_{\text{Hubs}}\) | \(K^*C\) | all Net | \(G_{FC}\) | \(G_{\text{Hubs}}\) | \(K^*C\) | all Net | \(G_{FC}\) | \(G_{\text{Hubs}}\) | \(K^*C\) | all Net | \(G_{FC}\) | \(G_{\text{Hubs}}\) | \(K^*C\) | all Net |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8.4 \times 10^8 | 5.9 \times 10^8 | 1050 | 1.05 \times 10^9 | 0.94 | 0.12 | 11.59 | 75 | 9.8 \times 10^5 | 6.0 \times 10^5 | 1088 | 1.1 \times 10^8 | 0.96 | 0.15 | 11.84 | 26 | 1.3 \times 10^6 | 7.7 \times 10^5 | 1153 | 1.9 \times 10^8 | 1 | 0.77 | 12.44 | 13 | 5.7 \times 10^5 | 7.4 \times 10^5 | 1200 | 3.0 \times 10^8 | 1 | 1 | 12.69 | 2 | 7.2 \times 10^5 | 6.2 \times 10^5 | 1090 | 1.4 \times 10^8 | 0.89 | 0.11 | 11.19 | 9 | 4.8 \times 10^5 | 3.8 \times 10^5 | 856 | 4.87 \times 10^7 | 0.87 | 0.165 | 7.94 | 1303 |
| 7.5 \times 10^7 | 5.5 \times 10^7 | 1007 | 3.6 \times 10^7 | 0.88 | 0.18 | 9.52 | 229 | 8 \times 10^5 | 5.6 \times 10^5 | 1014 | 6.4 \times 10^7 | 0.88 | 0.26 | 9.01 | 85 | 1.5 \times 10^6 | 6.9 \times 10^5 | 1124 | 1.17 \times 10^8 | 0.83 | 0.41 | 13.21 | 19 | 1.5 \times 10^6 | 6.8 \times 10^5 | 1135 | 7.3 \times 10^7 | 0.71 | 0.14 | 13.41 | 9 | 9.4 \times 10^5 | 6.03 \times 10^5 | 1014 | 6.7 \times 10^7 | 0.88 | 0.33 | 9.92 | 76 | 4.7 \times 10^5 | 3.7 \times 10^5 | 842 | 4.4 \times 10^7 | 0.87 | 0.15 | 7.7 | 1601 |
| 7.4 \times 10^6 | 5.4 \times 10^6 | 1022 | 7.6 \times 10^6 | 0.90 | 0.16 | 10.83 | 184 | 9.2 \times 10^5 | 5.9 \times 10^5 | 1106 | 1.1 \times 10^8 | 0.94 | 0.19 | 11.78 | 31 | 9.6 \times 10^5 | 6.4 \times 10^5 | 1043 | 1.3 \times 10^8 | 0.80 | 0.33 | 11.60 | 17 | 5.7 \times 10^5 | 7.4 \times 10^5 | 1199 | 3.0 \times 10^8 | 1 | 1 | 12.69 | 2 | 7.2 \times 10^5 | 6.2 \times 10^5 | 1090 | 1.4 \times 10^8 | 0.93 | 0.25 | 11.19 | 9 | 4.8 \times 10^5 | 3.8 \times 10^5 | 868 | 4.5 \times 10^7 | 0.89 | 0.1 | 7.86 | 1661 |
| 7.1 \times 10^5 | 5.1 \times 10^5 | 966 | 8.1 \times 10^7 | 0.88 | 0.2 | 10.82 | 188 | 7.1 \times 10^5 | 5.4 \times 10^5 | 1017 | 10^8 | 0.93 | 0.34 | 10.97 | 41 | 5.1 \times 10^5 | 6.5 \times 10^5 | 1018 | 1.2 \times 10^8 | 1 | 0.5 | 11.82 | 16 | 3 \times 10^5 | 5.9 \times 10^5 | 981 | 4 \times 10^7 | 1 | 0.5 | 11.27 | 2 | 8 \times 10^5 | 5.3 \times 10^5 | 926 | 9.0 \times 10^7 | 0.87 | 0.26 | 11.28 | 69 | 4.3 \times 10^5 | 3.5 \times 10^5 | 831 | 4.2 \times 10^7 | 0.84 | 0.17 | 7.64 | 1728 |
| 6.2 \times 10^5 | 4.48 \times 10^5 | 890 | 5.3 \times 10^7 | 0.88 | 0.22 | 8.50 | 662 | 6.3 \times 10^5 | 4.48 \times 10^5 | 884 | 5.5 \times 10^7 | 0.89 | 0.24 | 8.51 | 483 | 10^6 | 6.3 \times 10^5 | 1038 | 7.6 \times 10^7 | 0.94 | 0.5 | 11.0 | 26 | 1.2 \times 10^6 | 6.7 \times 10^5 | 1031 | 6.2 \times 10^7 | 0.93 | 0.42 | 10.44 | 14 | 8.2 \times 10^5 | 5.6 \times 10^5 | 968 | 7.5 \times 10^7 | 0.89 | 0.39 | 9.12 | 127 | 6.8 \times 10^5 | 5.1 \times 10^5 | 1046 | 6.3 \times 10^7 | 0.86 | 0.18 | 8.60 | 424 |
| 7.6 \times 10^5 | 5.3 \times 10^5 | 1048 | 6 \times 10^7 | 0.88 | 0.17 | 7.60 | 171 | 1.2 \times 10^6 | 6.9 \times 10^5 | 1139 | 9.9 \times 10^7 | 0.81 | 0.3 | 12.0 | 19 | 1.4 \times 10^6 | 6.7 \times 10^5 | 1148 | 7.8 \times 10^7 | 0.6 | 0.0 | 12.29 | 5 | 10^6 | 6.1 \times 10^5 | 1075 | 6.8 \times 10^7 | 0.88 | 0.25 | 9.44 | 83 | 6.8 \times 10^5 | 5.1 \times 10^5 | 861 | 4.4 \times 10^7 | 0.88 | 0.13 | 8.90 | 749 |
| 5.2 \times 10^5 | 3.9 \times 10^5 | 861 | 4.4 \times 10^7 | 0.88 | 0.13 | 8.90 | 749 | 5 \times 10^5 | 3.7 \times 10^5 | 824 | 3.5 \times 10^7 | 0.89 | 0.15 | 9.05 | 176 | 5.6 \times 10^5 | 5.4 \times 10^5 | 990 | 1.2 \times 10^8 | 0.94 | 0.41 | 11.70 | 33 | 5.6 \times 10^5 | 5.4 \times 10^5 | 986 | 1.2 \times 10^8 | 1 | 0.4 | 11.25 | 5 | 8 \times 10^5 | 5.5 \times 10^5 | 1007 | 9.4 \times 10^7 | 0.92 | 0.20 | 11.53 | 61 | 4.3 \times 10^5 | 3.5 \times 10^5 | 841 | 3.96 \times 10^7 | 0.87 | 0.12 | 7.51 | 2059 |