



Rich-cores in networks.

Ma, A; Mondragón, RJ

© 2015 Ma, Mondragón

Creative Commons Attribution License

For additional information about this publication click this link.

<http://qmro.qmul.ac.uk/xmlui/handle/123456789/15053>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

RESEARCH ARTICLE

Rich-Cores in Networks

Athen Ma*, Raúl J. Mondragón

Queen Mary University of London, School of Electronic Engineering and Computer Science, Mile End Road, London E1 4NS, United Kingdom

* athen.ma@qmul.ac.uk

Abstract

A core comprises of a group of central and densely connected nodes which governs the overall behaviour of a network. It is recognised as one of the key meso-scale structures in complex networks. Profiling this meso-scale structure currently relies on a limited number of methods which are often complex and parameter dependent or require a null model. As a result, scalability issues are likely to arise when dealing with very large networks together with the need for subjective adjustment of parameters. The notion of a *rich-club* describes nodes which are essentially the hub of a network, as they play a dominating role in structural and functional properties. The definition of a *rich-club* naturally emphasises high degree nodes and divides a network into two subgroups. Here, we develop a method to characterise a rich-core in networks by theoretically coupling the underlying principle of a rich-club with the escape time of a random walker. The method is fast, scalable to large networks and completely parameter free. In particular, we show that the evolution of the core in World Trade and *C. elegans* networks correspond to responses to historical events and key stages in their physical development, respectively.



OPEN ACCESS

Citation: Ma A, Mondragón RJ (2015) Rich-Cores in Networks. PLoS ONE 10(3): e0119678. doi:10.1371/journal.pone.0119678

Academic Editor: Bin Jiang, University of Gävle, SWEDEN

Received: July 8, 2014

Accepted: January 15, 2015

Published: March 23, 2015

Copyright: © 2015 Ma, Mondragón. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Full reference for each data has been given in the manuscript and all the data used are publicly available in numerous locations. Here is a list of locations where the authors obtained the data from: Zachary Karate club, Astrophysics collaborations (Astro), Condensed Matter collaborations (CondMat) *C.elegans*, American College football (football), Internet from <http://www-personal.umich.edu/~mejn/netdata/> Californian (CA) roads, Amazon recommendation (Amazon) from <http://snap.stanford.edu/data/US> airports (airports), Power grid (Power) from <http://toreopsahl.com/datasets/Protein> from <http://www3.nd.edu/~networks/resources.htm>

Introduction

Many networks exhibit a core/periphery structure which is important for the understanding of network properties and dynamics (for a review see ref. [1]). The constitution of such a structure often refers to a simple two-class partition [2–4], and a core is said to be comprised of densely inter-connected high-degree nodes which govern flows and impact adaptability, flexibility, and controllability in networks [5, 6]. The definition of core/periphery was formalised in [2] and numerous profiling methods developed were based on optimising a suitable fitness function, such as using a coreness value to define the density of links inside the core [2]; referring to a quality index with respect to the size of the expected core and the fuzziness of the boundary [7]; or applying Markov chains to describe random walks so as to index the coreness of individual nodes [8]. These methods rely on subjective fine-tuning due to the presence of one or more *free parameters* which are often obtained arbitrarily or by techniques such as simulated-annealing. Other examples include maximising the closeness centrality within the core by using an ensemble of random networks to define a coefficient that characterises the core [3], and this imposes a requirement of a statistical null model. Generally, these methods tend to be relatively

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

complex in nature and therefore scalability issues are likely to be encountered when applied to very large networks.

The notion of a rich-club is used to describe the connectivity between high degree nodes, and it has been applied to profile meso-scale properties in networks by examining the density of connections between high degree nodes [9–14]. A rich-club influences the functionality of a network, as demonstrated in the transmission of rumours in social networks [15], the delivery of information in the Internet [9], its strong effect on both the network assortativity and transitivity [11], and the organisation of the human connectome in neurodevelopment [16, 17]. Interestingly, the presence of a rich-club naturally divides a network into two parts; this means that the definition of a rich-club coincides with two fundamental prerequisites of a core/periphery structure: *high degree nodes* and a *two-class partition*. However, there is, at present, no general method to define which high degree nodes are members of the club [11, 13, 18].

Here, we present a *rich-core* method which profiles the core/periphery structure. The method brings together the concept of a rich-club and the diffusion of a random walker and defines a rich-core by examining the persistence probability of random walks among high degree nodes. We apply the method to a wide range of networks, ranging from man-made, social and biological, with a variety of network sizes. Furthermore, our results show that the core closely reflects the development and re-alignments of relationships over time in evolving networks. For example, changes in the core coincide with timing of historical events in trade development and with key developmental stages in a biological process. Anomalous nodes in the core can be uncovered with reference to null models by discriminating statistical differences. In addition, as the method does not levy any restriction on connectivity, the core can exhibit different properties, such as a core with an internal structure or a multi-core. The rich-core method is simple, fast, entirely free of any external parameters and applicable to very large networks, providing an effective and yet generic way to examine meso-scale properties in complex networks.

Results

Consider an unweighted and undirected graph. We rank the importance of the nodes in descending order of their degree, such that the node with the highest degree is ranked first and so on, and nodes are then re-labelled according to their rank. For a given node, we divide its links into two groups: those with nodes of a higher rank and those with a lower rank. More formally, a node with a rank r has degree k_r ; the number of links it shares with nodes of a higher rank is k_r^+ and the number of links with nodes of a lower rank is $k_r - k_r^+$. Core nodes are high degree nodes that are densely connected with each other [2], and we assume that the connectivity of a highly ranked node with other *higher* ranked nodes contributes towards the constitution of a core. Similarly, if a node has very few links with higher ranked nodes, it is likely to be a member of the periphery. To detect a core we propose this straightforward procedure (see [Materials and Methods](#) for the full technical description). Starting from the node with the highest rank, as r increases the number of links k_r^+ that node r shares with nodes of a higher rank fluctuates. There will be a node r^* where k_r^+ has reached its maximum, and from that node onwards k_r^+ is always less than $k_{r^*}^+$. This change in the connectivity among the highly ranked nodes defines the boundary of a (rich) core; the nodes with a rank less than or equal to r^* are the core and the rest belong to the periphery.

This pragmatic way of defining a core is related to the concept of a random walker in a network [8]. Consider a network which is partitioned into two sets: S_c is the core and S_p is the periphery. A random walker jumps from one node to another following a link between any pair of nodes, and the probability of the walker to visit a given node is proportional to the node's degree. The time it takes a random walker to escape from S_c to S_p is τ_c . If S_c is the set of nodes

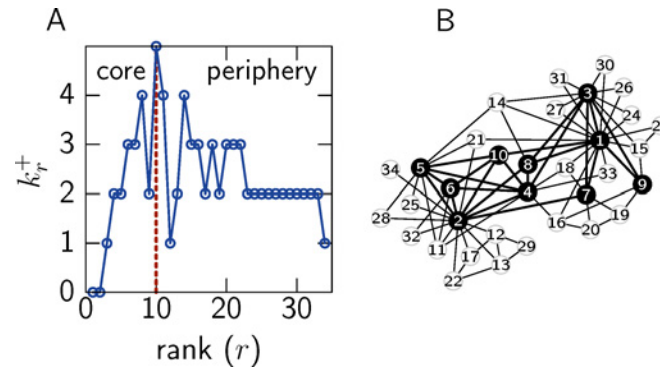


Fig 1. The rich-core for the Zachary Karate Club network. (A) The number of links k_r^+ that node r shares with nodes of a higher rank. The boundary of the core is marked by k_r^+ where k_r^+ is at its maximum, as indicated by the dotted line. (B) A graphical representation of the network with core (black) and periphery (white) nodes derived from the rich-core method.

doi:10.1371/journal.pone.0119678.g001

with the highest rank, as we include another node into this set the escape time will naturally increase. Eventually, S_c will contain all the nodes in the network, and the escape time will converge as the random walker is always contained in this set and has nowhere to escape to. This means that if we begin by putting only the top ranked node in set S_c and gradually increase the size of S_c by adding nodes in decreasing order of their rank, the escape time will always increase. If we consider the rate of change of the escape time as nodes are added to the set, the boundary of a rich-core is defined as the point in which the rate of increase in the escape time changes from a slow to a fast pace. This point coincides with the rank r^* defined above (as shown in [Materials and Methods](#)) which establishes a direct relationship between the cohesiveness of a core reflected by the escape time of a random walker and the boundary of a rich-core. Hence, the method itself does not impose any external parameter when defining a core. We examine the Zachary Karate Club network [19] which describes the friendships among members of the club. [Fig. 1A](#) shows that nodes are ranked in descending order of degree and the way in which k_r^+ changes as additional nodes are included. The rich-core of the network is bounded by the maximum k_r^+ and is formed by the 10 highest ranked nodes as shown in [Fig. 1B](#), including instructor Mr Hi and President John A., represented by node 1 and 2 respectively, who disagreed on the issue of lesson fees and led to a split in the club, and a number of their high degree followers.

Relative core size

The size of a core is an important property of a network as it has been suggested that a sizeable core makes a network more flexible and adaptable to changes [5], and a small core makes a network more controllable [6]. Here, we study the relative core size, c , which is the ratio between the number of nodes in the core N_c relative to the total number of nodes in the network N , across a wide range of networks. A network that has no periphery will have a relative core of 1, e.g. a fully connected network; a star network with N nodes has a relative core of $1/N$ and a core-less network has no links. [Fig. 2](#) shows the cores observed in networks which are different in size and structure (see Table A in [S1 File](#) for details), in part reflecting their functionality. For instance, the Amazon.com recommendation network and the Internet both have a relatively small core. The former is found to be disjoint and, as the network contains information about product recommendations, the results provide evidence of efficient information transfer

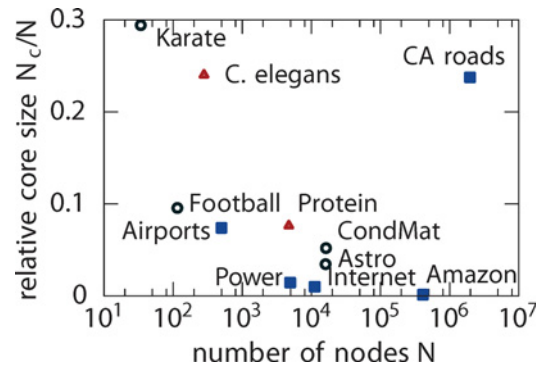


Fig 2. Relative size of the core in different kinds of real networks. Man-made (square): US airports (Airports) [22], Amazon.com recommendation (Amazon) [23], Californian (CA) roads [21], Internet [24] and Power grid (Power) [20]. Social (circle): Astrophysics collaborations (Astro) [25], Condensed Matter collaborations (CondMat) [25], American College football (Football) [26] and the Zachary Karate club (Karate) [19]. Biological (triangle): *C. elegans* [20] and Protein [27].

doi:10.1371/journal.pone.0119678.g002

within the network but only restricted to localised parts. The latter has a well connected single core which reflects its design for efficient routing between Autonomous System domains. The *C. elegans* neuronal network [20] has a relatively large single core, and this perhaps reflects the adaptability of the neuronal network to living conditions. The Californian road network [21] has a relatively large disjoint core which represents the existence of many crossroads, providing great flexibility in route choices as they present many possibilities between different geographical points. We did not observe any characteristic size of the core related to the origin of networks, i.e. man-made, social or biological. We compare our results with those obtained from using the core profiling (CP) method in [8] (Table A in S1 File). While the two methods identify many common high degree nodes in the core, the latter is more likely to define much larger cores by incorporating relatively low degree nodes (Figure A in S1 File) that do not necessarily fall under the formal definition of a core node [2].

Weighted and directed networks

The definition of a rich-core can firstly be extended to weighted networks [28, 29]. Consider w_{min} is the minimal weight linking two nodes in a network and the link between nodes i and j has a weight of w_{ij} . This link is represented by $\lceil w_{ij}/w_{min} \rceil$ links and the ranking is performed in units of the minimal weight. Each node is assigned to $\sigma_i = \sum_j \lceil w_{ij}/w_{min} \rceil$ links, and part of this quantity arisen from the node's linkage to nodes of a higher rank is referred to as σ_r^+ ; similarly, the remaining proportion, $\sigma_r - \sigma_r^+$, is the normalised weight that node r shares with nodes of a lower rank. The core boundary is node r^* such that $\sigma_{r^*}^+ > \sigma_r^+$ for $r > r^*$.

Similarly, the notion of a rich-core can equally be applied to directed networks by dividing links into in- and out- links and quantifying their corresponding weights. In this case the definition of a rich-core does not only depend on the weight of the nodes but also the direction of their links. An example is the assessment of web pages using PageRank where it has been observed that the popularity of a node is closely related to its in-degree [30]. Here, the direction of interest defines the in-links which determine the ranking (see Materials and Methods). Let $\sigma_r^{+in} + \sigma_r^{+out}$ be the total strength of interactions between node r and the nodes $r < r$, as both in- and out- links contribute towards the cohesiveness of a core, and the core boundary is the node r^* such that $\sigma_{r^*}^{+in} + \sigma_{r^*}^{+out} > \sigma_r^{+in} + \sigma_r^{+out}$ for $r > r^*$.

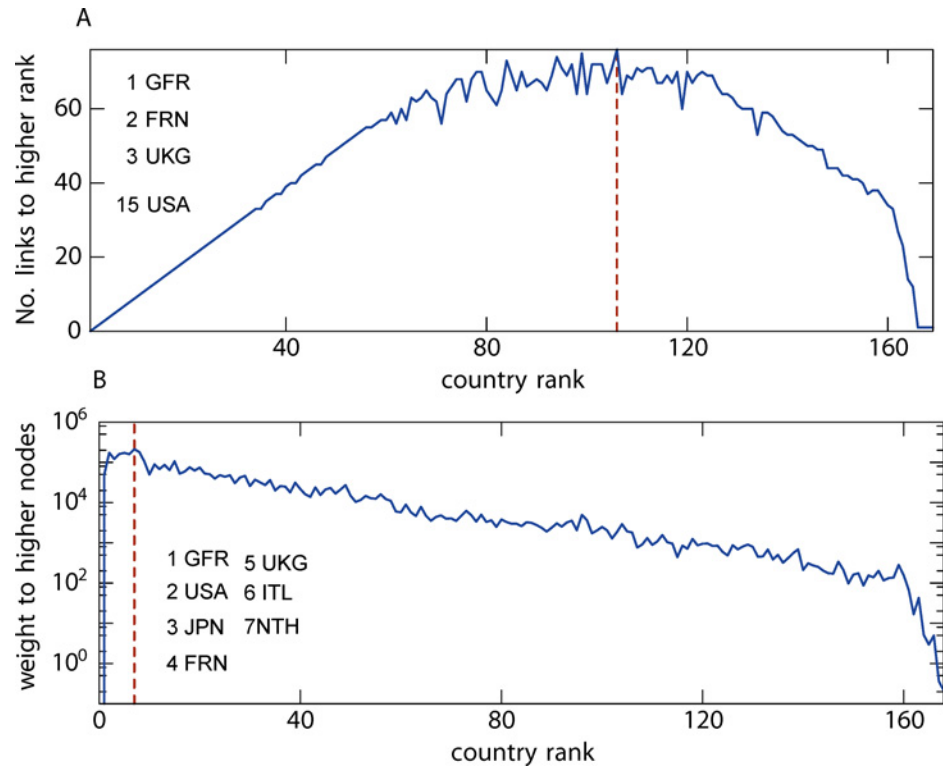


Fig 3. The World Trade network in 1990. (A) The unweighted undirected network which represents trade relationships between the countries. The core has 106 countries. (B) The weighted directed network representing the exchange of wealth between the countries. The 7 countries in core are listed.

doi:10.1371/journal.pone.0119678.g003

We refer to the World Trade network [31] as both an unweighted and weighted (directed) network whereby nodes are countries and links are trade channels; the latter can represent the direction of trade to specify an import or export relationship in a directed graph. The associated financial value can be seen as the weight of a given link. The overall connectivity of the network is found to be high as countries are interrelated in many ways. We first examine the network as an *undirected and unweighted* network and a link simply refers to the presence of trade. In 1990, a total of 60% of all the countries in the world were part of this network as a result of globalisation of trading at the time [32]. By ranking countries in descending order of their degree, Fig. 3A shows the number of trade relationships each of these countries has with countries of a higher rank. There are 106 countries in the core, which is similar in size to previous findings [8, 33]. The highest ranked nodes in the core (Germany, France, the UK) confirmed close trading relationships found among the countries within the European Union (EU) and the rest of the world. They were closely followed by other long established countries in the EU (e.g. Italy and the Netherlands) and the USA. Historically, these countries have well established trade with many other countries. The core provides an indication of the magnitude of interlinkage among developed countries and countries with strong manufacturing or agriculture (e.g. Brazil and Kenya); and countries outside the core are mostly confined to developing countries in Africa or countries that are very small in physical size (e.g. Andorra and San Marino). Now, if we take both the direction and weight into consideration, and nodes are ranked in the descending order of their export which is interpreted as the generated income towards

the country of origin, defining the weighted in-degree, and [Fig. 3B](#) shows that there are 7 members found in the core and they are consistent with the top exporters in the world at the time (in 1990) [[33](#)]. Coincidentally, these countries were also the top importers as there were high degrees of symmetry in financial values in the two trading directions [[32](#)].

Evolution

Networks are often found to be temporal in nature as they are subject to formation, dissolution and rewiring of links [[34](#), [35](#)]. Continuing with the example of World Trade we examine the evolution of the core between 1948 to 2000. While the number of participating countries continued to grow over time, the core of the directed and weighted network consists of a very selective group, corresponding to 4% to 6% of all the countries ([S1 Fig](#)). This can be explained by referring to the way in which World Trade has grown since the Second World War. International trading is said to be growing steadily but unevenly since the 1940s, as trade barriers were imposed by events such as the Cold War. The network was also strongly influenced by other key historical events, geographical distance, composition (e.g. products and services) and the nature of trade [[32](#)]. Throughout the 1980s and 1990s, there was a substantial reduction in the cost of shipping due to the explosion of air freight, the collapse of the Soviet Union leading to many independent countries, and the industrialisation of developing countries; all these events has shaped the development of trade worldwide, leading to a great leap in globalisation. This is closely reflected by the way the membership of the core has changed over time in [Fig. 4A](#). The USA, Germany, Japan, France and the United Kingdom were the top importers and exporters in the world during the period of study and it can be seen that these countries have been members of the core during the entire time. Canada was drifting in and out of the top ten in the World Trade ranking during the same period, and we can see a similar variability in its core membership. In addition, the economic reform in China, which started in the late 1970s, has led to a steady growth of $\sim 9\%$ in World Trade per year, and our results illustrate that China became a member of the core in 1997 which is just before China joining the World Trade Organisation in 2001.

Another example of how changes in the core tie in with key events in real networks can be found in the physical development of *C. elegans* neuronal connections in [Fig. 4B](#). The core of a fully developed worm contains 61 ([S1 Table](#)) out of a total of 302 neurones, and almost the entire core is developed within the first 500 minutes [[13](#)]. The formation of the core coincides with Embryogenesis, and it has been suggested that the highly connected neurones appear in the early development so as to minimise the energy cost by creating the core connections among key nodes that are not physically far apart [[36](#), [37](#)]; these connections can then be extended during the process of body elongation. New neurones are found after hatching in the late L1 larval stage at approximately 1250 minutes and the total number of neurones continues to grow until the start of the L4 larval stage at approximately 2400 minutes [[38](#)]. The post-hatching development causes the relative connectivity among the existing core neurones to decrease, resulting in a reduction in the overall size of the core, from its maximum size of 79 neurones at 1459 minutes to 61 neurones. The shrinking of the core coincides with the timing of Gonadogenesis.

Properties of the core

Certain properties of a core can be revealed by examining how the value of k_r^+ changes as additional nodes are included. For example, [Fig. 5A](#) shows k_r^+ plotted against r for the *C. elegans* network, and it can be seen that the top thirteen neurones are well connected, forming a tight cluster but sharing only one link with neurone RIAR which is ranked 14. Previously, only the

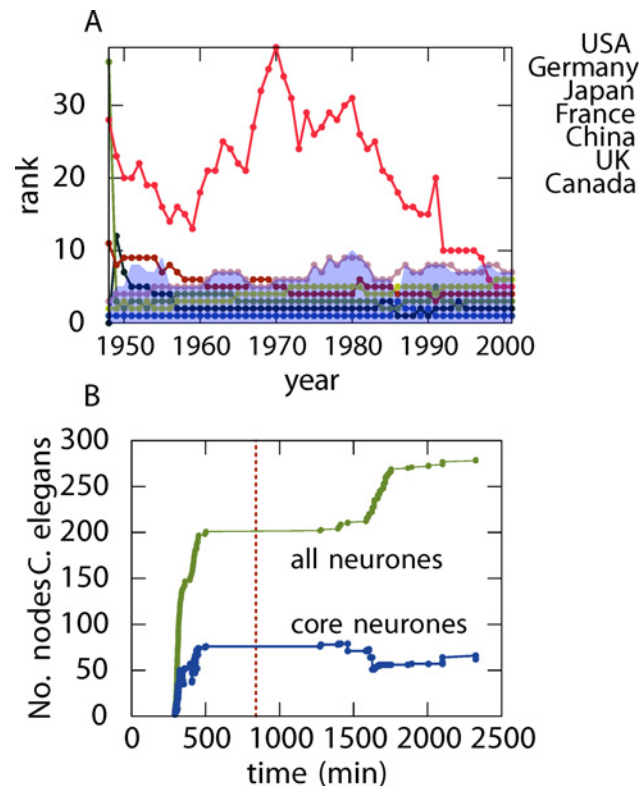


Fig 4. Evolution of the core of the World Trade and the *C. elegans* networks. (A) The World Trade network in which the shaded area in the background is the core region and countries in the core in a given year lie within this area. The countries shown are the members of the core in year 2000, and the individual lines are their ranks in a given year. The USA, Germany, Japan, France and the UK have already been found in the core, though there are small variations in the rank over the years. Canada is mostly found at the edge of the core during the period shown. China lies relatively far away from the core prior to the 1990s and it becomes part of the core in 1997. (B) Evolution of the core size for *C. elegans*. The numbers of nodes in the core and in the neural network are plotted against development time and the red dotted line marks the hatching time.

doi:10.1371/journal.pone.0119678.g004

top 14 nodes have been identified as the rich-club of the network [13]. Here, we show that the core is comprised of a set of closely connected high degree nodes interlinking with another set of high degree nodes immediately after neurone RIAR until the boundary of the core is reached. While the method itself does not require any null model to define a core, comparisons with a null counterpart, however, do provide a way to detect any anomalies with respect to the core size and connectivity among its members. We employ a randomisation method to create an ensemble of networks which in turn are used as a reference null model for comparison purposes. The randomisation is restricted to preserve the ranking of the nodes, that is the weight-/degree of the nodes (Materials and Methods). Continuing with the example of the *C. elegans* network, we create an ensemble of 100 networks with the same degree distribution as the original network. From the ensemble we evaluate the average number of links that node r has with nodes of a higher rank, i.e. $\langle k_r^+ \rangle$, and the standard deviation of this quantity. Fig. 5A also shows the number of links between a node of rank r and a node of rank $r' < r$ and the average number of links obtained from the null-model. The boundary of the core/periphery is the dotted blue line. The red line is $\langle k_r^+ \rangle$ and the pink shaded area demarcates two standard deviations from the mean value (numerically verified as the 95th percentile). This implies that nodes outside

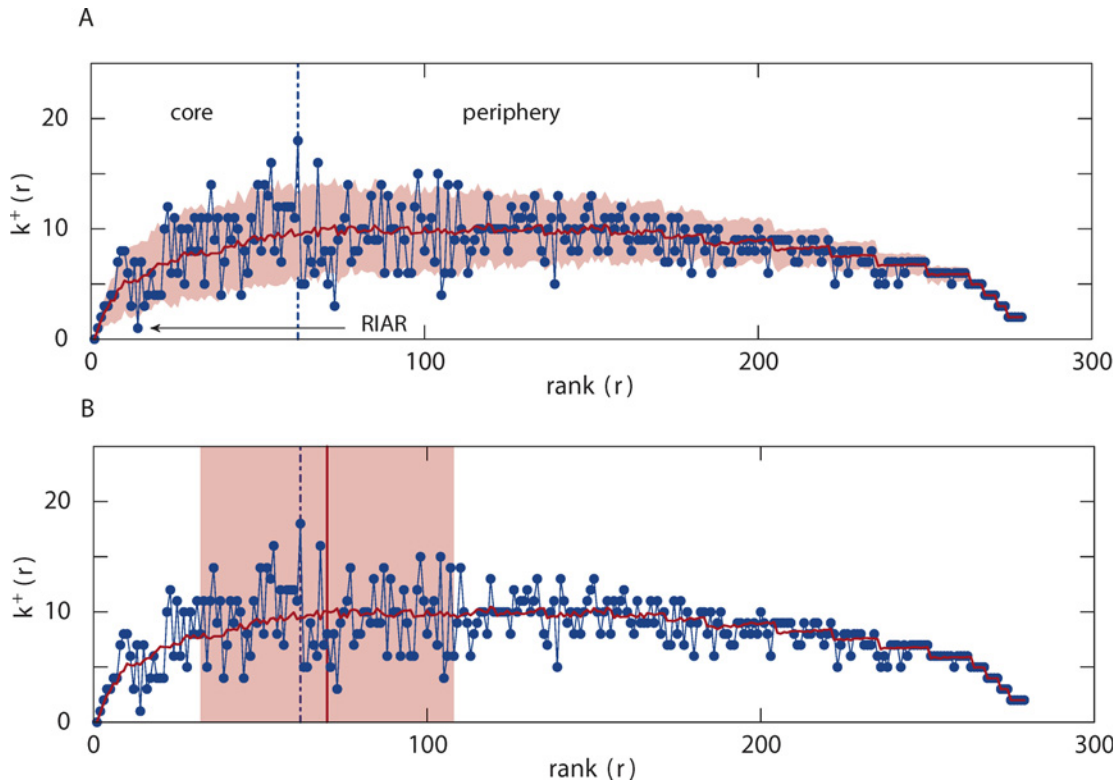


Fig 5. Comparison on the connectivity of the core between the empirical data and the null model for *C. elegans*. (A) The shaded area marks two standard deviations from the sample mean. Neurone RIAR is found to have an anomalous connectivity with the highest ranked nodes. (B) Comparison on the core size between the empirical data (blue line) and the null model (red line). Similarly, the shaded area shows two standard deviations.

doi:10.1371/journal.pone.0119678.g005

the shaded area can be considered anomalous. The fact that neurone RIAR shares only one link with nodes of a higher rank has been highlighted here, suggesting there is an anomaly in the connectivity. The randomisation can also be used to decide if the size of a core is within the expected value. Fig. 5B shows the average size of the core obtained from the ensemble of networks (red dotted line) and the pink area corresponds to two standard deviations from this mean. The dotted blue line is the size of the core obtained from the empirical data; as this value falls inside the pink area we can conclude that the size of the core for the *C. elegans* is what we expected.

Discussion

We develop a method to profile core/periphery in networks by coupling the notion of a rich-club and the persistence probability of random walks, resulting in the definition of a rich-core. In contrast to existing techniques that depend on one or more free parameters [2, 7, 8] or a null reference [3], our method only refers to the degree sequence and requires no prior knowledge on the network. As a result, the method is simple, fast and scalable, and particularly advantageous for profiling very large networks; as shown in the Amazon.com recommendation and Californian Road networks in which the method in [8] was found to be too computationally costly to run (Table A in S1 File). Our method is generic and is applicable to a wide range of networks as demonstrated by the results, and by examining the relative core size it is possible to gain an insight into the flexibility and controllability [5, 6] of a variety of networks. Though,

it is worth noting that approximating the persistence probability by a continuous function when profiling a core may lead to a small degree of deviation from the actual boundary, and this would particularly be the case among networks without a sharp core and periphery transition [7]. Similarly, degeneracy in the ordering of nodes originating from nodes with the same degree may cause the boundary to shift. In both cases, the impact on the actual core is found to be minimal.

Furthermore, we uncover characteristics of real networks through the profiling of the core/periphery structure, contributing towards the understanding of the significance of such meso-scale property. In particular, development and realignment of relationships are often found in networks with temporal nature. The core identified in the World Trade network represents the extent of trading among the developed countries and countries with a strong presence of manufacturing or agriculture, and the evolution of the core reflects closely how key historical events have shaped the patterns of import and export. Our results have demonstrated that the world's dominance in trading not only conformed to long-established countries but was also joined by emergent countries such as China in recent decades. Similarly, the *C. elegans* network shows that temporal changes of the core closely tie in with the key stages in the species' physical development. The proposed method does not impose any conditions on the connectivity inside the core and provides an effective way to reveal any internal organisation of a core, such as a multi-core. We show an effective way to identify any anomalies in the core by comparing the membership of the core of a given network with its null counterparts, and how this approach can be used to evaluate the expected size of a given core.

Compartmentalisation in networks has been suggested to be a key factor in facilitating network robustness [39]. A core organisation in networks is associated with structural properties such as symmetry, assortativity and hierarchy [1], and can be seen as an integrative function that provides redundancy for fluxes. However, we have yet to fully assess the impact of a core on network fragility, and the related dynamical properties are largely unexplored. A better understanding on these aspects will gain an insight into the functional role of a core and its impact on the overall network behaviour. We envision that the scope for cross-fertilisation here is enormous as networks, particularly those that are biological in nature, are often comprised of functional groups [40] and factors that may contribute towards their stability are currently still under debate.

Materials and Methods

Rich-club Coefficient and Core/Periphery Profile

To find the boundary of a rich-core, we examine the escape time it takes a random walker to leave a core. The escape time is related to the notion of persistence probability α which indicates the cohesiveness in a subgraph [8, 41]. The persistence probability of cluster S_c is

$$\alpha_c = \frac{\sum_{i,j \in S_c} \pi_i m_{ij}}{\sum_{i \in S_c} \pi_i} \quad (1)$$

where π_i is the probability that a random walker is in node i , and m_{ij} is the probability that a random walker moves from node i to node j . The escape time is $\tau_c = (1 - \alpha_c)^{-1}$. Assuming that α_x can be approximated by a continuous function $\alpha(x) = g(x)/f(x)$ where x is a continuous rank, the proposed method seeks to find the boundary of the core by locating the point at which the transition from a low to a high persistence probability accelerates. The function $\alpha(x)$ increases with x as the number of nodes in the cluster rises, and eventually converges to 1 when

all the nodes are included. Therefore, the first derivative of $\alpha(x)$ is always positive and the value of x for which the rate of increase is maximal is obtained when the second derivative is zero, i.e. $\alpha''(x) = 0$, where $\alpha''(x) = g''(x)/f(x) + 2g(x)f'(x)^2/f(x)^3 - 2f'(x)g'(x)/f(x)^2 - g(x)f''(x)/f(x)^2$. To first approximation $\alpha''(x^*) \simeq 0$ if $g''(x^*) = 0$ as $f(x)$ is a positive increasing function of x .

For undirected networks, α_c is given by the sum of the number of links between the nodes in S_c divided by the sum of the degrees of the nodes in S_c . If a_{ij} are the elements of the adjacency matrix, then

$$\alpha_c = \left(\sum_{i,j \in S_c} a_{i,j} \right) / \sum_{i \in S_c} k_i = 2 \frac{\sum_{i=1}^c k_i^+}{\sum_{i=1}^c k_i} \tag{2}$$

Again, assuming that α_c can be approximated with a continuous function $\alpha(x) = g(x)/f(x)$ where $g(x) = \int_1^x k^+(y)dy$ then $g''(x^*) = 0$ means that $k^+(x^*)$ has a maximum or a minimum at the value x^* , in this case we are interested in the maximum. We refer to the point x^* where $g''(x^*) = 0$ as the boundary of the rich-core and nodes in S_{x^*} are the members of the core.

The rich-club coefficient measures the density of links among (high degree) nodes and is defined as [9]:

$$\phi(r) = \frac{2E(r)}{r(r-1)} = \frac{2 \sum_{i=1}^r k_i^+}{r(r-1)} \tag{3}$$

where $E(r)$ is the number of links between the r nodes. The explicit relationship between the escape time and the rich-club coefficient is obtained by substituting Eq. (3) into Eq. (2), giving

$$\alpha_r = \frac{r(r-1)\phi(r)}{\sum_{i=1}^r k_i} \tag{4}$$

For weighted networks, as we are considering the weights as undirected multilinks, the same argument applies when defining a core. For directed networks, the persistence probability, α_c , is given by the ratio between the number of times a random walker transits inside the core and the number of times it visits the core. The former is proportional to the total number of links inside the core, which we denoted as $\sigma_{r^*}^{+in} + \sigma_{r^*}^{+out}$, as both in- and out- links provide paths for the random walker to move *within* core and hence contribute towards the persistence probability. The latter is given by the sum $\sum_{i \in S_c} \pi_i$. The in-degree k_i of node i is assumed to be a good approximation of π_i [30] which is characterised by the direction of interest (as demonstrated in the example of World Trade); hence, the number of times a random walker visits the core increases with r , as the nodes are ranked in decreasing order of their in-degree.

To find the core of a given network,

1. rank nodes in decreasing order of their weight (specifically, degree, in-degree and weight for undirected, directed and weighted networks respectively).
2. evaluate the number of links k^+ between node with rank r and nodes with rank $r' < r$
3. find the boundary of the core, defined by the node r^* where $k_{r^*}^+ > k_r^+$ for all $r > r^*$

Rank degeneracy

As numerous nodes can have the same degree, the ranking of nodes with equal degree is not strictly defined. This degeneracy in the ranking scheme would affect the determination of the boundary nodes, and hence the size of a core. To evaluate the effect of the degeneracy in the definition of a core we randomly re-rank the nodes with equal degree and measure the change

of the core nodes. We observe that this re-ranking only has a minor effect when defining a core.

Construction of Null Models

The null model is generated using the Zlatic *et al.* approach [29] which is a generalisation of Maslov, Sneppen and Zaliznyak method (MSZ) [42] to generate null models. Zlatic *et al.* redistributes the weights of the links by preserving the strength of the nodes as follows. If w_{min} is the minimum value of the weights in the original network then the rewiring is done by changing the weights of two pairs of links by an amount w_{min} . The rewiring consists of selecting two links at random and exchanging one of the end nodes of the first link with an end node of the second link.

Supporting Information

S1 File. Comparisons between the core profiling (CP) and rich-core (RC) methods.
(PDF)

S1 Fig. The relative core size of the directed and weighted World Trade network between 1948 and 2000.
(PDF)

S1 Table. Neurones that form the rich-core of the *C. elegans*. Rank, degree and birth time for all core neurones are shown.
(PDF)

Acknowledgments

We thank V. Latora, P. Vertes, J. Schormans and Chris Phillips for their comments on the manuscript, X. Lu and C. Gray for early discussions related to this research. We also thank the reviewers for their insightful and invaluable comments.

Author Contributions

Conceived and designed the experiments: AM RJM. Performed the experiments: AM RJM. Analyzed the data: AM RJM. Contributed reagents/materials/analysis tools: AM RJM. Wrote the paper: AM RJM.

References

1. Csermely P, London A, Wu LY, Uzzi B. Structure and dynamics of core/periphery networks. *Journal of Complex Networks*. 2013; 1:93–123. ArXiv preprint arXiv:1309.6928 doi: [10.1093/comnet/cnt016](https://doi.org/10.1093/comnet/cnt016)
2. Borgatti SP, Everett MG. Models of core/periphery structures. *Social networks*. 1999; 21(4):375–395. doi: [10.1016/S0378-8733\(99\)00019-2](https://doi.org/10.1016/S0378-8733(99)00019-2)
3. Holme P. Core-periphery organization of complex networks. *Physical Review E*. 2005; 72(4):046111. doi: [10.1103/PhysRevE.72.046111](https://doi.org/10.1103/PhysRevE.72.046111)
4. Da Silva MR, Ma H, Zeng AP. Centrality, network capacity, and modularity as parameters to analyze the core-periphery structure in metabolic networks. *Proceedings of the IEEE*. 2008; 96(8):1411–1420. doi: [10.1109/JPROC.2008.925418](https://doi.org/10.1109/JPROC.2008.925418)
5. Csete M, Doyle J. Bow ties, metabolism and disease. *TRENDS in Biotechnology*. 2004; 22(9):446–450. doi: [10.1016/j.tibtech.2004.07.007](https://doi.org/10.1016/j.tibtech.2004.07.007) PMID: [15331224](https://pubmed.ncbi.nlm.nih.gov/15331224/)
6. Liu YY, Slotine JJ, Barabasi AL. Controllability of complex networks. *Nature*. 2011; 473 (7346):167–173. doi: [10.1038/nature10011](https://doi.org/10.1038/nature10011) PMID: [21562557](https://pubmed.ncbi.nlm.nih.gov/21562557/)
7. Rombach MP, Porter MA, Fowler JH, Mucha PJ. Core-periphery structure in networks. *SIAM Journal on Applied mathematics*. 2014; 74(1):167–190. doi: [10.1137/120881683](https://doi.org/10.1137/120881683)

8. Della Rossa F, Dercole F, Piccardi C. Profiling core-periphery network structure by random walkers. *Scientific reports*. 2013; 3.
9. Zhou S, Mondragón RJ. Accurately modeling the Internet topology. *Physical Review E*. 2004; 70(6):066108. doi: [10.1103/PhysRevE.70.066108](https://doi.org/10.1103/PhysRevE.70.066108)
10. Colizza V, Flammini A, Serrano MA, Vespignani A. Detecting rich-club ordering in complex networks. *Nature*. 2006; 2(2):110–115.
11. Xu XK, Zhang J, Small M. Rich-club connectivity dominates assortativity and transitivity of complex networks. *Physical Review E*. 2010 Oct; 82(4):046117. doi: [10.1103/PhysRevE.82.046117](https://doi.org/10.1103/PhysRevE.82.046117)
12. van den Heuvel MP, Sporns O. Rich-Club organization of the human connectome. *The Journal of Neuroscience*. 2011; 31(44):15775–15786. doi: [10.1523/JNEUROSCI.3539-11.2011](https://doi.org/10.1523/JNEUROSCI.3539-11.2011) PMID: [22049421](https://pubmed.ncbi.nlm.nih.gov/22049421/)
13. Towilson EK, Vértes PE, Ahnert SE, Schafer WR, Bullmore ET. The rich club of the *C. elegans* neuronal connectome. *The Journal of Neuroscience*. 2013; 33(15):6380–6387. doi: [10.1523/JNEUROSCI.3784-12.2013](https://doi.org/10.1523/JNEUROSCI.3784-12.2013) PMID: [23575836](https://pubmed.ncbi.nlm.nih.gov/23575836/)
14. Mondragón RJ. Network null-model based on maximal entropy and the rich-club. *Journal of Complex Networks*. 2014;.
15. Masuda N, Konno N. VIP-club phenomenon: Emergence of elites and masterminds in social networks. *Social Networks*. 2006; 28(4):297–309. doi: [10.1016/j.socnet.2005.07.005](https://doi.org/10.1016/j.socnet.2005.07.005)
16. Ball G, Aljabar P, Zebari S, Tusor N, Arichi T, Merchant N, et al. Rich-club organization of the newborn human brain. *Proceedings of the National Academy of Sciences*. 2014;p. 201324118.
17. Grayson DS, Ray S, Carpenter S, Iyer S, Dias TGC, Stevens C, et al. Structural and functional rich club organization of the brain in children and adults. *PloS one*. 2014; 9(2):e88297. doi: [10.1371/journal.pone.0088297](https://doi.org/10.1371/journal.pone.0088297) PMID: [24505468](https://pubmed.ncbi.nlm.nih.gov/24505468/)
18. Valverde S, Solé RV. Self-organization versus hierarchy in open-source social networks. *Physical Review E*. 2007; 76(4):046118. doi: [10.1103/PhysRevE.76.046118](https://doi.org/10.1103/PhysRevE.76.046118)
19. Zachary WW. An information flow model for conflict and fission in small groups. *Journal of anthropological research*. 1977;p. 452–473.
20. Watts DJ, Strogatz SH. Collective dynamics of small-world networks. *Nature*. 1998; 393(6684):440–442. doi: [10.1038/30918](https://doi.org/10.1038/30918) PMID: [9623998](https://pubmed.ncbi.nlm.nih.gov/9623998/)
21. Leskovec J, Lang KJ, Dasgupta A, Mahoney MW. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*. 2009; 6(1):29–123. doi: [10.1080/15427951.2009.10129177](https://doi.org/10.1080/15427951.2009.10129177)
22. Colizza V, Pastor-Satorras R, Vespignani A. Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics*. 2007; 3(4):276–282. doi: [10.1038/nphys560](https://doi.org/10.1038/nphys560)
23. Leskovec J, Adamic LA, Huberman BA. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*. 2007; 1(1):5. doi: [10.1145/1232722.1232727](https://doi.org/10.1145/1232722.1232727)
24. Chen Q, Chang H, Govindan R, Jamin S. The origin of power laws in Internet topologies revisited. In: *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*. Proceedings. IEEE. vol. 2. IEEE; 2002. p. 608–617.
25. Newman ME. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*. 2001; 98(2):404–409. doi: [10.1073/pnas.98.2.404](https://doi.org/10.1073/pnas.98.2.404)
26. Girvan M, Newman ME. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*. 2002; 99(12):7821–7826. doi: [10.1073/pnas.122653799](https://doi.org/10.1073/pnas.122653799)
27. Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature*. 2001; 411(6833):41–42. doi: [10.1038/35075138](https://doi.org/10.1038/35075138) PMID: [11333967](https://pubmed.ncbi.nlm.nih.gov/11333967/)
28. Serrano MA. Rich-club vs rich-multipolarization phenomena in weighted networks. *Physical Review E*. 2008; 78(2):026101. doi: [10.1103/PhysRevE.78.026101](https://doi.org/10.1103/PhysRevE.78.026101)
29. Zlatic V, Bianconi G, Díaz-Guilera A, Garlaschelli D, Rao F, Caldarelli G. On the rich-club effect in dense and weighted networks. *The European Physical Journal B*. 2009; 67(3):271–275. doi: [10.1140/epjb/e2009-00007-9](https://doi.org/10.1140/epjb/e2009-00007-9)
30. Fortunato S, Boguñá M, Flammini A, Menczer F. Approximating PageRank from in-degree. *Lecture Notes in Computer Science*. 2008; 4936:59. doi: [10.1007/978-3-540-78808-9_6](https://doi.org/10.1007/978-3-540-78808-9_6)
31. Gleditsch KS. Expanded trade and GDP data. *Journal of Conflict Resolution*. 2002; 46(5):712–724. doi: [10.1177/0022002702046005006](https://doi.org/10.1177/0022002702046005006)
32. World Trade Report 2013. World Trade Organisation. 2013; Available: http://www.wto.org/english/ress_e/publications_e/wtr13_e.htm. Accessed 30 October 2013.
33. Fagiolo G, Reyes J, Schiavo S. The evolution of the world trade web: a weighted-network analysis. *Journal of Evolutionary Economics*. 2010; 20(4):479–514. doi: [10.1007/s00191-009-0160-x](https://doi.org/10.1007/s00191-009-0160-x)

34. Leskovec J. Dynamics of large networks. ProQuest; 2008.
35. Kostakos V. Temporal graphs. *Physica A: Statistical Mechanics and its Applications*. 2009; 388(6):1007–1023. doi: [10.1016/j.physa.2008.11.021](https://doi.org/10.1016/j.physa.2008.11.021)
36. Bullmore E, Sporns O. The economy of brain network organization. *Nature Reviews Neuroscience*. 2012; 13(5):336–349. PMID: [22498897](https://pubmed.ncbi.nlm.nih.gov/22498897/)
37. Nicosia V, Vértes PE, Schafer WR, Latora V, Bullmore ET. Phase transition in the economically modeled growth of a cellular nervous system. *Proceedings of the National Academy of Sciences*. 2013; 110(19):7880–7885. doi: [10.1073/pnas.1300753110](https://doi.org/10.1073/pnas.1300753110)
38. Byerly L, Cassada R, Russell R. The life cycle of the nematode *Caenorhabditis elegans*. I. Wild-type growth and reproduction. *Developmental Biology*. 1976; 51(1):23. doi: [10.1016/0012-1606\(76\)90120-2](https://doi.org/10.1016/0012-1606(76)90120-2) PMID: [988845](https://pubmed.ncbi.nlm.nih.gov/988845/)
39. May RM. Will a large complex system be stable? *Nature*. 1972; 238:413–414. doi: [10.1038/238413a0](https://doi.org/10.1038/238413a0) PMID: [4559589](https://pubmed.ncbi.nlm.nih.gov/4559589/)
40. Melián CJ, Bascompte J. Food web cohesion. *Ecology*. 2004; 85(2):352–358. doi: [10.1890/02-0638](https://doi.org/10.1890/02-0638)
41. Piccardi C. Finding and testing network communities by lumped Markov chains. *PloS One*. 2011; 6(11): e27028. doi: [10.1371/journal.pone.0027028](https://doi.org/10.1371/journal.pone.0027028) PMID: [22073245](https://pubmed.ncbi.nlm.nih.gov/22073245/)
42. Maslov S, Sneppen K, Zaliznyak A. Detection of topological patterns in complex networks: correlation profile of the internet. *Physica A: Statistical Mechanics and its Applications*. 2004; 333:529–540. doi: [10.1016/j.physa.2003.06.002](https://doi.org/10.1016/j.physa.2003.06.002)