A Game Theory Inspired Approach to Stable Core Decomposition on Weighted Networks

Srinka Basu and Ujjwal Maulik, *Senior Member, IEEE*

Abstract—Meso-scale structural analysis, like core decomposition has uncovered groups of nodes that play important roles in the underlying complex systems. The existing core decomposition approaches generally focus on node properties like degree and strength. The node centric approaches can only capture a limited information about the local neighborhood topology. In the present work, we propose a group density based core analysis approach that overcome the drawbacks of the node centric approaches. The proposed algorithmic approach focuses on weight density, cohesiveness, and stability of a substructure. The method also assigns an unique score to every node that rank the nodes based on their degree of core-ness. To determine the correctness of the proposed method, we propose a synthetic benchmark with planted core structure. A performance test on the null model is carried out using a weighted lattice without core structures. We further test the stability of the approach against random noise. The experimental results prove the superiority of our algorithm over the state-of-the-arts. We finally analyze the core structures of several popular weighted network models and real life weighted networks. The experimental results reveal important node ranking and hierarchical organization of the complex networks, which give us better insight about the underlying systems.

Index Terms—Weighted core, stable core, core structures, game theory, voting mechanism, complex network analysis, weighted network

1 Introduction

Networks are often used to model complex systems from various domains where the system's entities are represented by nodes, while the relationship between entities are represented by edges [1]. Structural analysis based on such network representation has revealed numerous insights about the complex systems.

Much of the recent work on the structural analysis of the networks has focused on the node-centric measures of local topology, such as, vertex degree, strength, clustering coefficient [1]. Network structures, however, can be described using a meso-scale or intermediate scale and a global scale perspective as well. In particular, identification of meso-scale network structures makes it possible to discover features that might not be apparent either at the local level of nodes and edges or at the global summary level.

Core analysis, where a core indicates a densely connected substructure, is one such meso-scale network analysis problem. Core structure analysis of network has received rather less attention from the scientific community in network science. The study of core structures, however, is an important problem. The core nodes in a network often play a different role from the rests of the network. As a result, core structure analysis has found its applicability in versatile domains. For example, in biological domain, the core analysis helps in the study of brain networks [2], [3] while in protein interaction networks the core structure helps analysis and prediction of protein functions [4], [5]. In social and information networks the core structures help to predict the most influencing spreader [6], detect the elite structure [7], study and evaluate the degree of cooperation [8]. In scientific collaboration networks core analysis helps the prediction of possible collaborators [9], [10]. Core structure analysis is found useful in the risk analysis in the financial networks [11], in the analysis of the Internet graph [12], in the control of the traffic flow in transportation networks [13], in identification of the key transmission lines in the power grid networks [14]. Distinguishing the core structures also helps in more informative visualization of the complex networks [15].

The existing approaches on core analysis, like k-core [16], s-core [17] are based on node degree and strength respectively. A limitation of the design of these approaches is that it fails to exploit both the information (degree and strength) simultaneously on a weighted graph. Thus producing unsatisfactory results especially on networks where core consists of nodes with low degree and high strength or high degree and low strength. Moreover, by design, these methods find only self-similar cores [12], [18]. Yet another limitation of these approaches is that these are also very sensitive to noise [19]. The real life networks, however, are heterogeneous in nature and prone to noise. The existing approaches tend to fail to discover the core structures accurately in such cases.

The insufficiencies of the existing approaches motivates to design new core analysis method with objectives - i) to identify the densely and strongly connected structures even when node degree and strength are uncorrelated, ii) to identify heterogeneous cores of various sizes and shapes, iii) to be stable against noise. A recent developed approach, KShell decomposition [20] that combines node degree and strength, closely satisfy the above objectives. However, KShell is again a node property based approach. In an attempt to meet the above objectives, we propose an

• *The authors are with the Department of Computer Science and Engineering, Jadavpur University, Kolkata, India. E-mail: srinka.basu@gmail.com, umaulik@cse.jdvu.ac.in.*

Manuscript received 22 June 2015; revised 5 Nov. 2015; accepted 29 Nov. 2015. Date of publication 22 Dec. 2015; date of current version 30 Mar. 2016. Recommended for acceptance by L.B. Holder.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TKDE.2015.2508817

algorithmic approach to the core analysis that is based on the group property. We also propose a method to rank the nodes based on their core-ness. Thus adding flexibility in the selection of the core sizes. The proposed framework borrows the game theoretic concept of voting mechanism. The group based method, over the node based methods of core analysis, is more informative about the topology of the neighborhood of a node as it also measures the importance of the nodes one is connected with. Moreover, the group based methods are found to be more stable against random noise. Finally, the merit of using the game-theoretic methods is that, in general, it considers the entities as rational and emphasizes on the self-organization of the entities formed based on some purposes, not for optimizing some global objectives [21].

To study and compare the performance of the proposed core decomposition method with the state-of-the-art we develop a new synthetic benchmark with planted core structure. The benchmark is an ensemble of weighted synthetic networks whose properties can be controlled by various parameters. We next study the performance of the core decomposition methods on a weighted lattice that is not expected to show any core structure. To test the robustness of our algorithm, we carry out the tests on noisy networks. The experimental results show that the proposed method performs significantly better than the state-of-the-art on the all the synthetic instances.

We finally apply the proposed core decomposition approach for the structural analysis of several popular weighted network models namely weighted exponential model (WE) [22], weighted preferential selection model (WPS) [23] and weighted stochastic block model (WSBM) [24] and weighted real life networks including character co-appearance networks, transportation network, scientific collaboration networks and online social network (OSN). The experimental results show that the proposed core analysis technique reveal important node ranking and hierarchical organization of the complex networks. The study, thus unveil the critical organizing components of the networks and help to gain better insight about the underlying system.

The organization of the paper is as follows. Section 3 introduces the notations used and the relevant definitions of graph theory. Section 4 formally defines the proposed core and elaborates the proposed game theoretic framework. Sections 5 and 6 discusses the experiments on synthetic benchmark networks and the real life networks. Finally, the scope is concluded in Section 7 with summary, conclusions and future work.

2 Related Work

An attempt to formally define core, has led its association to the graph theoretic model, clique. Cliques are completely connected subgraphs [25], [26], [27]. In real life networks, however, clique structures are rare as it imposes the restriction on complete connectivity of the structure. To this end, clique relaxation models, that relax the properties like connectivity, minimum degree are found to be more suitable. k-core, k-clique, k-plex are some of the popular clique relaxation models [28].

The concept of k-core [16] decomposes a network based on the node property of degree. The problem of k-core decomposition can be easily solved in polynomial time by recursively removing the vertices of degree less than k [29], [30]. As a result, the k-core model, over the other clique relaxation models, has gained significant popularity as a network analysis tool in various applications [2], [18], [31], [32], [33], [34], [35], [36], [37].

A limitation of k-core is that it works only on unweighted networks. Many real life networks include additional information on link weight that quantifies the strength of a tie. Considering such networks as unweighted graphs results in a significant loss of information. s-core is a generalized k-core approach for weighted network analysis [17]. The s-core is defined as the subgraph consisting of all nodes with node strength $s_i > s$, where s_i is the sum of the weights of the edges incident on a particular node i and s is a threshold value. $s-\mathrm{core}$ decomposition is based on iterative removal of edges with strength less than s. The common drawbacks of k-core and s-core are - i) they fail on networks where node degree and strength are uncorrelated, ii) find only self-similar cores [12], [18], iii) sensitive to random noise.

KShell decomposition [20] combines both the information to identify the cores in the network. The approach uses a multiplicative form to combine the node degree and node strength and quantizes the resultant score for the purpose of core decomposition. KShell can identify core structures when degree and strength are uncorrelated. KShell is also more stable against noise as it weaves together two different local topological measures.

A group-centric method to core analysis is another approach that overcome the limitations of the existing approaches. On unweighted graphs, γ-quasi-clique [28], which require the edge density of the induced subgraph to be at least γ, is one such method. γ quasi clique defines cohesiveness without requiring every node to have degree over the specified threshold, thus relaxing strict self-similarity constraint among the nodes of a structure. γ-quasi-clique being a group-density based approach is also more stable against noise.

Though the problem of finding maximum quasi clique is NP-Complete, the polynomial time approximation approach involves iterative removal of nodes until the γ quasi clique is found [38]. The heuristics for γ-quasi-clique based graph partition approaches are found in [39], [40], [41], [42]. These approaches are highly sensitive on the orders in which nodes are removed and yield very different results for different vertex ordering. Moreover, the existing works on γ-quasi-clique ignore the information about the edge strength.

Another related concept we study is the authority score [43]. Originally designed for directed networks, on undirected network the score ranks the nodes based on the notion whether a node is strongly connected with many highly ranked vertices. The authority score is determined by the principle eigen vector of $A^T A$, where A is the adjacency matrix of the network. Though, the authority score do not directly reveal the core structures, but the distribution of the authority scores can be used for the study of the core analysis. For example, by setting a cut-off and selecting all nodes having authority score above the threshold gives us a strongly and densely connected substructure representative of the core.

A few recent works [44], [45] have used a related concept of collective reputation for assessment and prediction of scientific contributions. [44] exploits fitness for the long-term predictability of citation pattern where fitness is a collective measure capturing the community's response to a scientific-work. In [45], an algorithmic approach is proposed that computes reputation of users in scientific online communities based on the quality of papers, reputation of users, and credit of authors.

3 Preliminaries

We assume a network is represented as a simple undirected weighted graph $G=(V,E,W)$ with a finite vertex set $V = \{1,\ldots,n\}$, a finite edge set $E \subseteq V \times V$ and a finite weight set $W = \{w_{ij}\}$ for every edge $e_{ij} \in E$. $w_{ij} = 0$ when nodes $i \mathrm{and} j$ are not connected. We assume all the edge weights are normalized so that the maximum possible edge weight is 1. $G(S)$ denote the subgraph induced by the vertices $S \subseteq V$. The weight (strength) of a vertex $v \in S$ within the subgraph is $w_{G(S)}(v) = \sum_{u \in G(S)} w_{uv}$. Weight (strength) of a subgraph $G(S)$ is the sum of the weights of all the vertices within $G(S)$ and is denoted by $w(G(S))$. The cross edge weight of a vertex $v \in S$ is $X_{G(S)}(v) = \sum_{u \notin S, e_{uv} \in E} w_{uv}$. For any vertex $v \in V$, the pth order open neighborhood $N_p(v)$ is the set of vertices connected to v by a path of length less than or equal to p. The closed pth order neighborhood of vertex v is $N_p[v] = N_p(v) \cup \{v\}$.

Definition 1. *The edge weight density of $G(S)$, $\delta(G(S))$ is the ratio of $w(G(S))$ to twice the maximum possible edge weight i.e., $|S|(|S|-1)$, where the weights are normalized and maximum weight is* 1,

$$\delta(G(S)) = \frac{w(G(S))}{|S|(|S|-1)}.$$

In a weighted graph G, a weighted γ quasi clique is defined below.

Definition 2. *A weighted γ quasi clique is a subset of vertices $\emptyset \neq S \subseteq V$ that induces a subgraph $G(S)$ such that $\delta(G(S)) \geq \gamma$ where $0 < \gamma \leq 1$.*

Given G and γ, the size of a quasi clique S, is the cardinality of S. A quasi clique is called maximal if it is not contained in a larger quasi clique and maximum if S has the largest cardinality in the graph.

Definition 3. *A subgraph $G(S)$, induced by $S \subseteq V$ is a locally maximal weighted γ quasi clique if $\delta(G(S)) \geq \gamma$, where $0 < \gamma \leq 1$ and addition of any node neighboring to S decreases the value of $\delta(G(S))$.*

In this paper we refer to weighted γ quasi clique as quasi clique, in short.

4 The Proposed Core Structure

The objective of the new core analysis methods are - i) to identify the strongly and densely connected structures irrespective of how the node degree and strength are correlated, ii) to allow cores of various shapes, iii) to be stable against random noise.

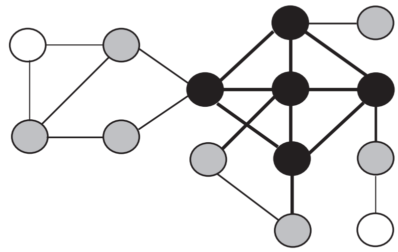

Fig. 1. (β, γ) Quasi Core decomposition. The black colored nodes form a $(1, 0.8)$ core. All the similar colored nodes form a shell.

4.1 The (β, γ) Core

The proposed framework to core problem borrows the graph-theoretic concept of weighted γ-quasi clique and the game theoretic concept of voting mechanism.

Definition 4. *Given a graph $G=(V,E,W)$, a nonempty set $S \subseteq V$ is a (β, γ) core, if and only if all the following conditions hold true without prioritizing any one.*

- *i) For every node $i \in S$, $G(S) \cap N_p[i]$ forms a locally maximal weighted γ quasi clique, where $0 < \gamma \leq 1$ and $p \geq 1$*
- *ii) probability of co-appearance of any two nodes in $G(S)$ is at least β, $0 < \beta \leq 1$*

We introduce another related concept of (β, γ) shell. A shell is defined as a set of vertices with maximum cardinality belonging to a (β, γ) core but not in any higher core.

Definition 5. *Given a graph $G=(V,E,W)$, a nonempty set $S \subseteq V$ is a (β, γ) shell, if and only if all the following conditions hold true.*

- *i) $S \subseteq (\beta, \gamma)$ core*
- *ii) $S \nsubseteq (\bar{\beta}, \bar{\gamma})$ core where $\bar{\beta} > \beta$ and $\bar{\gamma} > \gamma$*
- *iii) $\nexists \bar{S} \neq S$, such that $|\bar{S}| > |S|$ and $\bar{S}$ is a (β, γ) shell.*

The parameters β, γ determine the quality of a core in terms of stability and edge weight density. Higher the value of the parameters, the more stable and the more dense the structure is.

By definition, (β, γ) core follows a hierarchical structure. While the parameter γ has a strong influence on the weight density of the cores, β influences the stability and both the parameters allow to obtain a hierarchical structure of cores. Indeed, for a fixed value of γ, β_1-cores are included in β_2-cores if $\beta_1 > \beta_2$, i.e., β_1-cores are sub-cores of β_2-cores. Similarly, for a fixed value of β, γ_1 is a sub-core of γ_2 if $\gamma_1 > \gamma_2$. Also, (β_1, γ_1) core is a sub-core of (β_2, γ_2) core if $\beta_1 > \beta_2, \gamma_1 > \gamma_2$.

A (β, γ) core S is not necessarily a single connected component. Moreover, for a given graph G there might exist a vector (β, γ), for which the (β, γ) core is an empty set. On the other hand, there also might exist a vector (β, γ), for which the (β, γ) core is equal to V.

Example 1 demonstrate the (β, γ) core decomposition and its hierarchical organization.

Example 1. Fig. 1 shows a weighted undirected graph of 14 nodes and 20 edges. As shown in Fig. 1, the edge weights are proportional to the edge widths. Let us assume the darkest edges have weight 1 and the rests have weight 0.5. The (β, γ) core analysis reveals, the back colored nodes form a $(1, 0.8)$ core. The grey colored nodes along with the black ones form a lower core. All the same colored nodes form a (β, γ) shell.

4.2 The Algorithm

We formulate the core decomposition problem as a coalition formation game (*CFG*) on a given weighted graph $G = (V, E, W)$. N is the set of players in *CFG*, where player $i \in N$ represents node $i \in V$ and $|N| = |V|$. Before discussing the detail game play we define the following design parameters.

- Potential of a player i within a subgraph S is defined by $f(w_{G(S)}(i), X_{G(S)}(i))$ where $i \in S$. A node with higher value of $X_{G(S)}(i)$ and lower $w_{G(S)}(i)$ is classified as a low potential node. In our case, we use the function, $f(a, b) = a + Cb$ for some constant C. Empirically, we set the value of C as -1 which gives the best result.
- Collective worth of a subgraph S is defined by $g(S) = \delta(G(S))$. The chosen function for collective worth represents the weight density of the subgraph S.
- Marginal contribution of a player i in a subgraph S is defined by $\delta(G(S)) - \delta(G(S \setminus \{i\}))$ where $i \in S$. This capture how important a node is to a subgraph S.
- Scoring order is the lexical ordering of $N_p(i)$ that i follows when assigning the core scores. In the current work, after comparing several other orders, we consider the non-decreasing order of potential f of all the players in $N_p(i)$ as it gives the best performance.
- γ score strategy is the strategy node i follows while voting node $j \in S \subseteq N_p(i)$. γ score of j is the collective worth of the subgraph S where the potential of every node in S is at least as much as that of j. In other words it captures the maximum weight density that i can attain with j and other nodes who are as wealthy as j is. Once all the nodes assigns γ scores to their bounded neighbors, the γ scores of every node is averaged to obtain the final γ score.
- β score strategy, is followed by an arbiter and it controls the size and stability of the structure. β score is computed by counting and normalizing the number of votes one receive.

The scoring strategies adopted in the game determine the quality of the cores. The objective is not just to have high density cores but also the one that is more likely to appear and larger in size.

The game play is defined below.

- The game starts with a random player i.
- Every player i selects prospective coalition $C_i = N_p(i)]$.
- $\forall j \in C_i$, i selects a player $j \in C_i$ following the scoring order. If j has negative marginal contribution, j is eliminated from C_i. i assigns a γ score to j following the scoring strategies.
- Once all the players assign γ scores to their neighbors, the arbiter assign β score following the γ scoring strategy.

Once the game play is over, post processing steps are conducted to group the nodes or to assign a continuous ranking. Keeping the post processing steps separate, adds flexibility to our framework. We adopt three different post processing schemes as discussed below.

1) When β and γ are known, all the nodes which has their scores equal to or higher than the given (β, γ) vector are selected as core. However, knowing the exact value of β and γ is not a trivial problem. The determination of the appropriate values of β and γ for the core study is taken as a future scope of research.
2) To cluster the nodes based on the (β, γ) scores, we round off the scores and sample for every value of (β, γ) in $\{0, 0.01, \dots, 1\} \times \{0, 0.01, \dots, 1\}$ space.
3) Assign a continuous rank to the nodes based on the (β, γ) scores. We use a multiplicative approach to determine the Alternative Core Score (ACS).

4.3 Interpretation of Core Scores

The proposed framework provides various ways to use and interpret the assigned core score by selecting any of the post processing approaches and determine the final core structure. For example, one can select a core structure for a specific value of (β, γ), or can sample the core structure from the parameter space of β and γ, or can combine the information that yield a continuous node ranking. The post processing approach can be chosen based on the context of the problem. For example, in Section 5.1 when studying the synthetic benchmarks, as the size of the core is known, we select a given fraction of nodes with top ACS value. As another example, in Sections 7.1, 6.4 and 6.3 we study the distribution of the ACS value. While the shape of the distribution is proved to be most useful to analyze the null model in Section 7.1, it is the mean and the standard deviation of the distribution that supports the study of the transportation and the online social network systems in Sections 6.4 and 6.3 respectively. In the study of the scientific collaboration networks, we study and compare the performance of both (β, γ) core decomposition and ACS. Thus, with a minimum a priori knowledge the user can select an appropriate post processing policy. The above feature of the proposed framework gives additional flexibility and can be used beneficially for discrete or continuous classification of core.

4.4 Computational Complexity

The runtime complexity of the Algorithm 1 is dominated by the complexity of assigning the γ scores in steps $5 - 14$. In the worst case, every node assigns a γ score to all other $n - 1$ nodes, and thus, the worst case requiring $O(n)$ computations. For n nodes the total complexity becomes $O(n^2)$. The post processing steps performed require either constant or $O(n)$ computations. Thus, the overall worst case complexity is $O(n^2)$.

On a scale-free network, however, the complexity of the Algorithm 1 is much lower than $O(n^2)$, and in average case becomes $O(n \log(n))$, where the expected number of players to whom one vote is $\log(n)$.

Algorithm 1. Assigning Core Scores

Let $N \leftarrow G(V)$
Let w be the weight matrix of G
Let p be the co-appearance matrix of size $n \times n$
Let β, γ be the 0 vectors of dimension n
$Q \leftarrow \emptyset$
for all $i \in N$ **do**
  Let $C_i \leftarrow N_p(i)$
  Let $\pi_i \leftarrow$ the score order of i
  for all $j \in \pi_i$ **do**
    if Marginal contribution of j is negative **then**
      $C_i \leftarrow C_i \setminus \{j\}$
      $p[i,j] \leftarrow i$ assign γ score to j
    end if
  end for
end for
for all $i \in N$ **do**
  $\gamma[i] \leftarrow mean_j(p[i,j])$
end for
for all $i \in N$ **do**
  $\beta[i] \leftarrow count_j(p[i,j] \neq 0)$
  Normalize β
end for
For every node $i \in N$ prepare the list Q containing $(\beta[i], \gamma[i])$
Return Q

5 Model Evaluation

To ensure a systematic and thorough evaluation of the proposed model, we first examine the correctness of our method using an ensemble of networks with a planted core structure. We then test the performance on null models considering networks which do not have any meaningful core-structure. Finally to evaluate the robustness of the proposed method we carry out experiment on noisy network data.

5.1 Synthetic Network with Planted Core Structure

We develop a family of synthetic networks that have a single core structure. We use $CS(N, d, \tau_c^d, \tau_{nc}^d, \tau_c^w, \tau_{nc}^w, k_c, k_{nc}, \mu)$ to denote this ensemble of networks. Each network in the ensemble has N nodes, where dN of the nodes are core nodes, $(1-d)N$ of the nodes are non-core nodes and $d \in [0,1]$. The average degree of the core nodes is k_c and that of the non-core nodes is k_{nc}. The degree of core nodes and non-core nodes follow power law distribution with exponents and average as τ_c^d, k_c and τ_{nc}^d, k_{nc} respectively. The weights of core edges and non-core edges follow power law distribution with exponents as τ_c^w and τ_{nc}^w respectively. The mixing parameter μ controls the fraction of edges that connect the core nodes to the non-core nodes. We first build a graph with two connected components with random edge distribution, one representing the core structure (with parameters dN, k_c, τ_c^d and τ_c^w) and the other representing the non-core structure (with parameters $(1-d)N$, k_{nc}, τ_{nc}^d and τ_{nc}^w). Subsequently, we rewire the edges of the connected components with probability μ. The finally obtained graph contains a core structure when $\tau_c^d \geq \tau_{nc}^d, \tau_c^w \geq \tau_{nc}^w, k_c \geq k_{nc}$. This model can be generalized to build a hierarchical multi-core model. However, the simplistic single core model is sufficient for testing the correctness of the approaches.

We fix $N = 100, d = 1/2, \tau_c^d = 2.9, \tau_c^w = 2.9$ and generate 100 instances for each of the combination of the parameter values $\tau_{nc}^d = \{2.01, 2.1, 2.2, \ldots, 2.9\}$; $\tau_{nc}^w = \{2.01, 2.1, 2.2, \ldots, 2.9\}$; $\mu = \{0.1, 0.2, \ldots, 1\}$; $k_{nc}/k_c = \{0.1, 0.2, \ldots, 1\}$; correlation between degree and strength as $\{-1, -.9, \ldots, 1\}$; correlation between strength and clustering coefficient as $\{-1, -.9, \ldots, 1\}$.

5.1.1 Results and Observations

We use the metric sensitivity (the fraction of core nodes correctly identified) to compare the performance of the algorithms on the synthetic networks with planted core structure. We compare the proposed method with the performance of KShell, s-Core, k-Core, γ quasi clique decomposition (QC), weighted authority score (AR) in terms of sensitivity. For the proposed method we use the ACS and select the top 50 percent nodes for the experiments. For the AR as well, the top 50 percent nodes are considered as core nodes. The top core determined by KShell, s-Core and k-Core are considered for the evaluation. For QC, we sampled the value of γ from the domain $\{.1, .15, .2, \ldots, 1\}$ and considered the partition that results in maximum sensitivity.

Fig. 2 shows the performance of the algorithms averaged over 100 different instances with same parameter value. As shown in the Fig. 2a, when the weight exponent changes, the performance of k-core, QC remains unaffected, as both the approach ignores the edge weight information. However, the performance of the s-core and KShell drops significantly as the weight exponent of the non-core nodes approaches that of the core nodes. The sensitivity of our algorithm remains mostly unaltered, closely followed by that of AR. On change of degree exponent of the non-core nodes, in the Fig. 2b, the performance of k-core, QC drops significantly. Sensitivity of KShell also drops but its performance is better than that of k-core, QC. The sensitivity of our algorithm, AR, and s-core mostly remain unaltered, with the proposed method scoring highest sensitivity amongst all.

On change of mixing parameter, the inter-connectedness of the core structure changes and it becomes increasingly difficult to identify the core boundary clearly. As shown in the Fig. 2c the performance of all the algorithms drop significantly as μ approaches 1, except that for k-core. This is possibly because k-core analyses the network based on node degree and the orientation of edges do not affect its performance. It is important to note that even when $\mu = 1$, the algorithms still manages to find the core structure with sensitivity above 0.4, that is because the degree and strength of the core nodes still remain high in the network.

The Fig. 2d shows as the average degree (and strength) of the non-core nodes approaches that of the core nodes, the performance of all the algorithms get affected. Finally we observe the performance of the algorithms as the correlation between degree and strength and that between strength and clustering coefficient varies, shown in the Figs. 2e and 2f respectively. In almost all the instances our algorithm outperform all other algorithms, except for the high values of mixing parameters, where the k-core scores the highest.

5.2 Random Network without Core Structure

A good algorithm to find core structures should indicate a clear difference between a network with core structure and

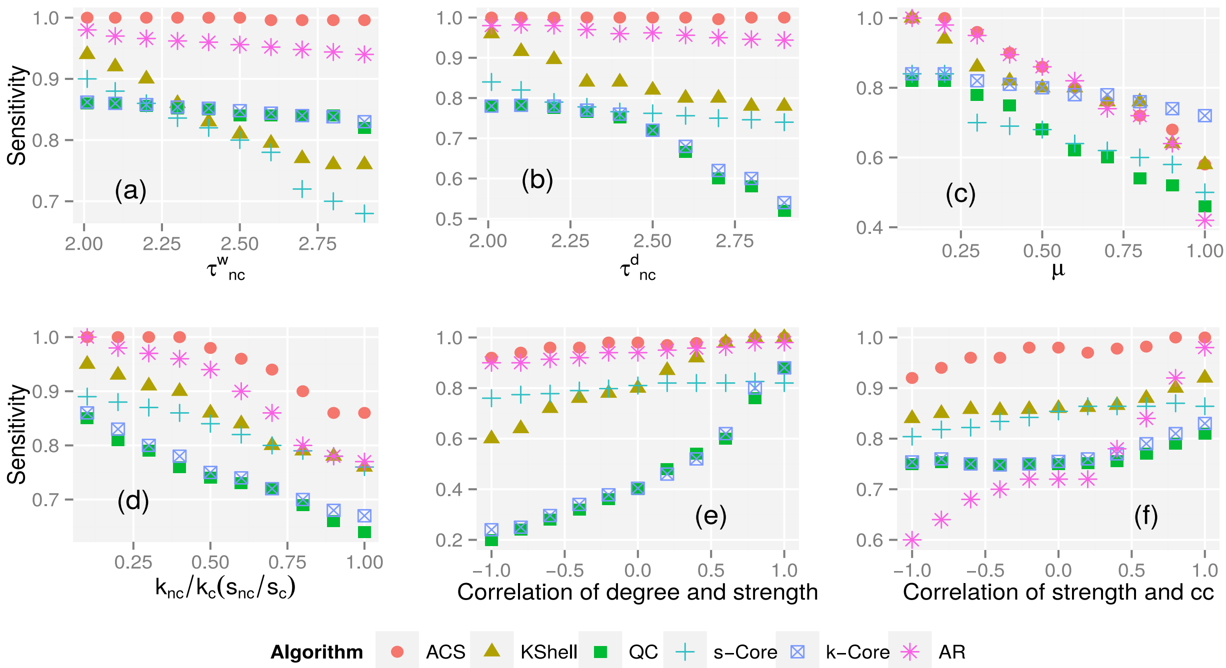

Fig. 2. Sensitivity of the core analysis algorithms on imposed core structure networks. (a) shows the change in sensitivity when τ^w_{nc} changes, while τ^w_c is kept constant at 2.9. (b) shows the change in sensitivity when τ^d_{nc} changes, while τ^d_c is kept constant at 2.9. (c) shows the variation in sensitivity when μ changes. (d) shows the change in sensitivity with change in k_{nc}/k_c and s_{nc}/s_c. (e) and (f) shows the variation in sensitivity when correlation between degree and strength and that between strength and clustering coefficient (cc) varies from -1 to 1.

a null model, that is, a network without any meaningful core structure. We consider a lattice, which does not exhibit any meaningful core structure as the null model. All nodes in a lattice have the same degree and all edges have same weight.

5.2.1 Results and Observations

The alternative core score of every node in a lattice converges to the same value. By taking averages over 100 runs, we observe that the alternative core score of each node is similar, and converge to equal core scores. Hence, our method correctly indicates that lattice networks have no meaningful core structure. This example illustrates that instead of just looking at the core score magnitude, one should examine the distribution of the core scores. This can be done visually, or by computing the variance.

5.3 Synthetic Networks with Random Noise

Presence of noise is a very common phenomenon when dealing with complex real life systems. Noisy data occur due to many reasons including but not limited to erroneous measurements, sampling bias, unresponsiveness to surveys. Thus the study of stability of an approach against noise is essential. We study the stability of the core structures when edge weights are perturbed with random noise. We follow the method of adding random noise to edge weights as mentioned in [46]. The noise added over an edge weight initially equal to w_{ij}, is equally distributed between $[-\alpha\, w_{ij}, \alpha\, w_{ij}]$, where $0 < \alpha < 1$ is the rate of perturbation. The synthetic benchmark described in Section 5.1 are used in the experiment with parameters as $N = 100, d = 0.5, \tau^d_{nc} = 2.1, \tau^d_c = 2.9$, $\tau^w_{nc} = 2.1$, $\tau^w_c = 2.9$, $k_{nc}/k_c = 0.2, s_{nc}/s_c = 0.2$, correlation between degree and strength and that between strength and cc as 0.25.

5.3.1 Results and Observations

We compare the cores of the perturbed networks with that of the original network by computing: i) the normalized mutual information (NMI) [47], ii) the fraction of nodes identified as part of core in perturbed networks but do not belong to the core of the original network (false positive) and iii) the fraction of nodes belong to the core of the original network but not identified as part of core in perturbed networks (false negative).

As k-core and QC works on unweighted networks, change in edge weight does not affect their performance. Thus, these two algorithms are not considered in the experiment. We compare the proposed method (nodes with top 50 percent ACS) with the performances of KShell (the top core), s-Core (the top core) and AR (the top 50 percent nodes). The Fig. 3 shows the performance when averaged over 100 instances for every value of $\alpha \in \{0, 0.1, \ldots, 0.5\}$. In the presence of random noise the proposed method outperforms other methods as it attains the maximum NMI, minimum false positive and minimum false negative with lowest variation for each metric. The performance of KShell closely follows the performance of our algorithm, while s-core is found to be the most vulnerable to noise.

6 Core Analysis of Complex Networks

We next study the core structures to mine popular weighted network models and weighted real life networks using the proposed core and compare how it is related to that revealed by KShell, s-Core, k-Core, QC and AR.

For QC, we first find the maximum γ quasi cliques by sampling γ from the domain $\{0.1, 0.15, 0.2, \ldots, 1\}$. We then generate a disjoint partition such that if a node belongs

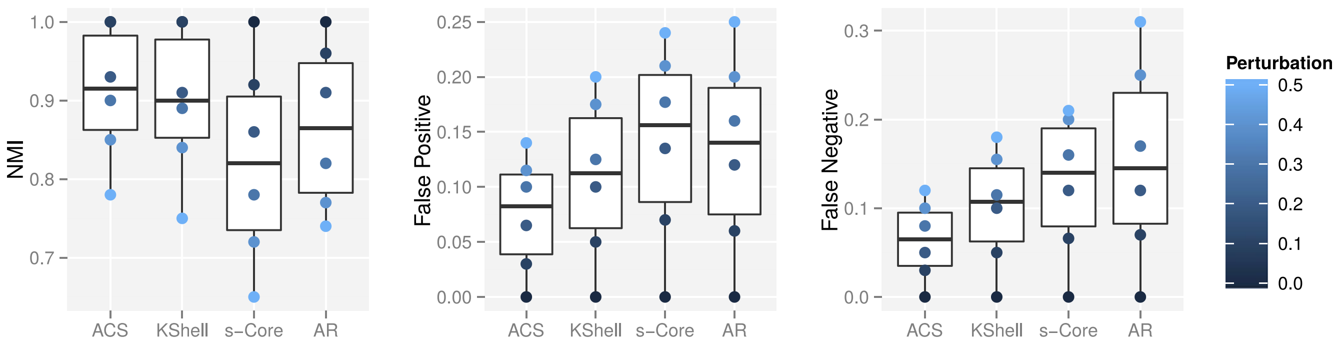

Fig. 3. Performance of the core analysis algorithms when the edge weights are perturbed with random noise. The figures, from left, shows the change in NMI, false positive and false negative, respectively, when the rate of perturbation varies from 0 to 0.5.

to multiple γ quasi cliques, the one with maximum γ value is considered.

6.1 Weighted Network Models

Network models help us generate networks that closely follows the properties of real life networks. For the generality of the study we first conduct the tests on three popular weighted network models - i) weighted exponential model [22] ii) weighted preferential selection model [23], iii) weighted stochastic block model [24]. In WE model weight of every new edge added is proportional to the degree of the incident node resulting in exponential degree and weight distribution [22]. WPS focuses on a strength driven preferential attachment in which new vertices connect more likely to vertices which are more central in terms of the strength of interactions. In WPS, addition of a fixed weight edge introduces variation to the existing weights across the network [23]. The degree and weight distributions follow power law in WPS model. WSBM generalizes the stochastic block model by annotating edges with weights drawn from an exponential family distribution. The parameters used for generating instances of the models are: i) number of nodes is 100, ii) a new node is attached to five existing nodes in preferential models, iii) the fraction of weight induced by a new edge in WPS is $\delta = 0.5$, iv) the number of groups of nodes in WSBM model is 2 with size 20 (core) and 80 (non-core), v) the edges in WSBM follow Bernoulli's distribution with probability $0.1, 0.09, 0.009$ for intra-core connectivity, core and non-core connectivity and intra-non-core connectivity respectively, v) the weights in WSBM follow normal distribution with parameters $(\mu = 100, \sigma = 1)$, $(\mu = 25, \sigma = 1)$, $(\mu = 4, \sigma = 1)$ for intra-core connectivity, core and non-core connectivity and intra-non-core connectivity, respectively. The parameters in WSBM are chosen in a manner such that the generated network contains a strongly connected dense group and a weakly connected sparse group. Table 1 shows the global summary statistics of the networks. The results are obtained by averaging on 100 instances of each network settings.

Fig. 5a shows the distribution of ACS on the network models. As shown in the figure a sharp drop in the ACS distribution and a visible right fat tail indicate the existence of a core structure. To validate further we examine the network layouts. As shown in Fig. S.1, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TKDE.2015.2508817 all the networks exhibit core structure. However, for WSBM the high strength nodes do not share a strongly connected neighborhood, unlike that of WE and WPS model. This can be attributed to the feature of the WSBM that allows us to control the weight and connection distribution between groups but do not allow to control the neighborhood connection pattern, resulting in high strength nodes with weakly connected neighborhood or low strength nodes with strongly connected neighborhood.

Fig. 6a shows the correlation between the ACS and other core detection methods. As shown by the bargraphs, the high similarity between ACS and other weighted core detection methods in WPS network is possibly because of the network property of high similarity between degree and strength. In WE and WPS no significant similarity is observed between ACS and other methods. This indicate that the hub (weighted) nodes in WPS also share a strongly connected neighborhood unlike other network models WE and WSBM.

6.2 Co-Appearance Networks

Social network analysis tools can help us understand and provide additional insights into the network of characters in various novels. Here we consider two co-appearance networks. The first (Lesmiserable) is a representation of the characters of the classic epic novel Lesmiserable authored by Victor Hugo, where the nodes represent the characters of the novel while the links represent the two corresponding characters co-appeared in the selected chapters of the book. The edges are weighted by the frequency of co-appearance of the characters [48]. Another co-appearance

TABLE 1
Data Set Description

Network	N	E	$<d>$	$<w>$	t_{glb}	r_{ds}	r_{sc}
WE	100	485	9.7	0.0526	0.173	0.78	0.396
WPS	100	485	9.7	1.98	0.386	0.995	−0.336
WSBM	100	388	7.7	31.36	0.095	0.665	−0.12
Lesmiserable	77	252	6.6	21.3	0.49	0.90	0.14
KJB	1.7K	9.1K	10.3	18.5	0.16	0.94	−0.34
OSN	1.9K	14K	14.6	39.2	0.11	0.91	−0.05
TFL	307	354	2.3	3.0	0.03	0.61	0.13
Geom	6.2K	12K	3.9	7.3	0.49	0.90	−0.035
HepTh	7.6K	16K	4.1	4.0	0.49	0.80	−0.18

The table shows the name of the networks, number of nodes (N), edges (E), average degree ($<d>$), average weight ($<w>$), global clustering coefficient (t_{glb}), correlation between degree and strength (r_{ds}), correlation between strength and clustering coefficient (r_{sc}) of the largest connected component.

TABLE 2
Characters with Top-10 ACS of Lesmiserable Network

Characters	ACS	Characters	ACS
Gavroche	1	Joly	0.62
Marius	0.98	Javert	0.56
Valjean	0.79	Eponine	0.54
Bahorel	0.67	Fantine	0.53
Mabeuf	0.64	Enjolras	0.52

network (KJB), collected in [49], [50] is constructed from the proper nouns (places and names) of the King James Bible. A node represents a proper noun and an edge indicates that two nouns appeared together in the same verse. The edge weights show how often the two nouns appeared together. Table 1 shows the global summary statistics of the networks.

In Lesmiserable network, the characters with top 10 ACS are mentioned in Table 2. It is important to note that though Valjean is the node with maximum degree, it is not the most important character according to the core score. This reveals that the characters with which Valjean co-appear are rarely related to each other, resulting in a low core value. On the contrary, the character Gavroche in spite of having lower degree and strength, result in the highest core score. This indicates the acquaintances of Gavroche not just know each other but also co-appear frequently. This is apparent from the Fig. S.2, available online, which shows the neighborhood of Gavroche and Valjean. More importantly, only those neighbors of Valjean, who are also connected to Gavroche are the most connected neighbors of Valjean. The figure also reveals, seven other top ACS scoring nodes that co-appear with Gavroche, namely, Marius, Valjean Mabeuf, Bahorel, Joly, Javert, Fantine, Enjolras are also strongly and densely connected amongst themselves.

In Fig. 4 we show the comparison amongst the different core decomposition approaches. It is worth noting that ACS method identifies Gavroche as a core node whereas s-core and Authority rank identifies Valjean as the top core. KShell, ranks Marius, Enjolras Combeferre, Feuilly, Courfeyrac, Bahorel, Bossuet, Joly equally with Gavroche. From the Fig. 4 it is also apparent that the majority of the characters have very low core score and do not participate in the central core structure. We attribute these characters as non-core or peripheral characters.

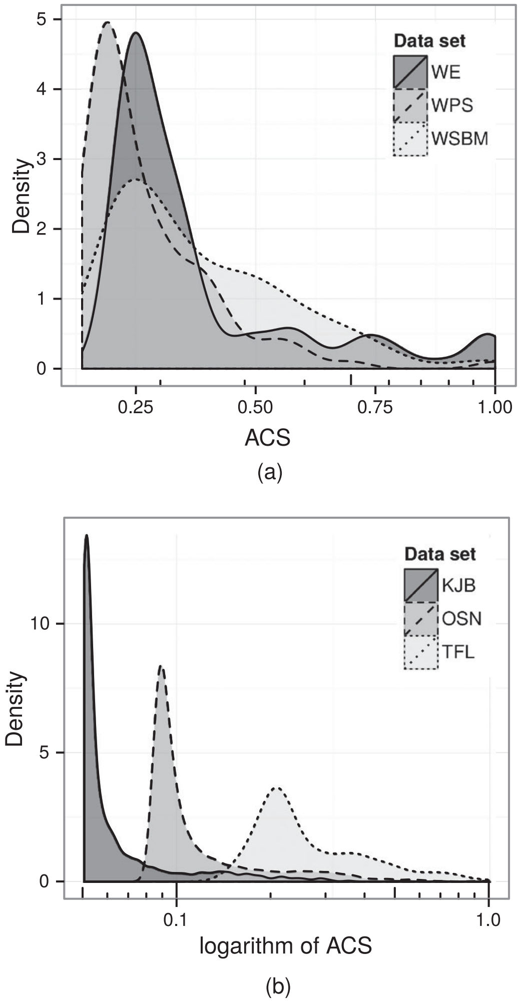

Fig. 5. Sub-figure (a) shows the distribution of ACS on WE, WPS, and WSBM networks and sub-figure (b) shows the distribution of logarithm of ACS on KJB, OSN, and TFL networks. (a) shows ACS in network models tend to follow a log-normal distribution while in (b) shows the ACS on real networks follows a power-law distribution. Different colors indicate different data sets.

Table 3 shows the top 15 ACS scoring proper nouns of KJB network. It is interesting to note that the names of places like Israel, Jerusalem, Egypt scores very high as they are well connected with other characters and places. Also, it is important to note that there is a sharp drop in ACS score between the first position(Israel) and the second position (Judah) nodes and the forth position (Jerusalem)

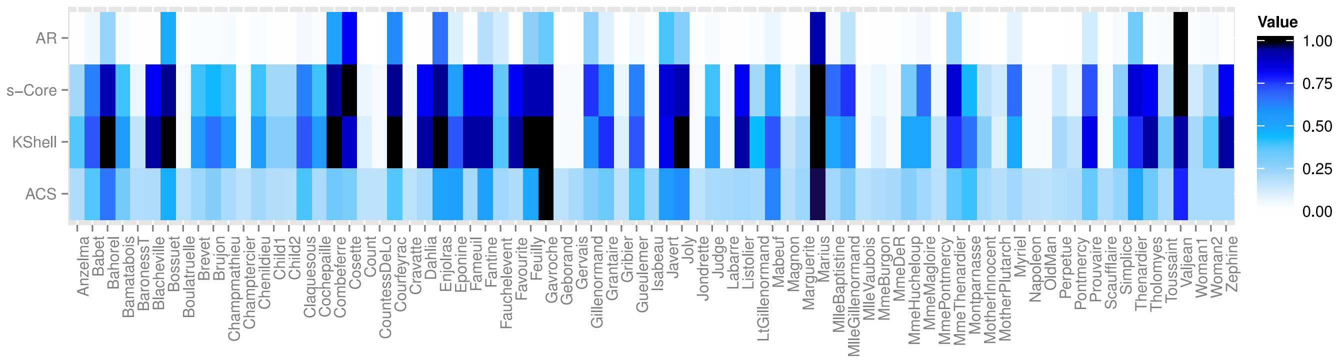

Fig. 4. Coreness of the characters of Lesmiserable as computed by the algorithms ACS, KShell, s-core, and AR.

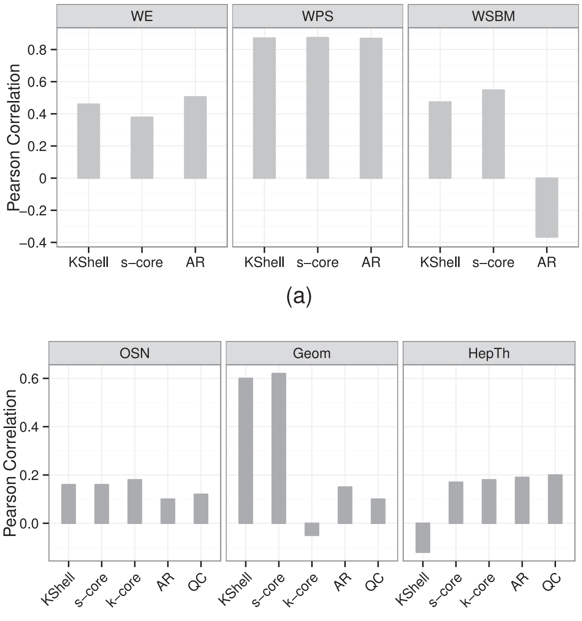

Fig. 6. Bargraphs shows the correlation of ACS with other core scores, namely, KShell, s-core, QC, k-core and AR on - (a) WE, WPS, and WSBM network model and (b) on OSN, Geom and HepTh real networks. The vertical axis shows the value of the Pearson correlation while the horizontal axis indicate the core detection methods for each data set.

and the fifth position (Ephraim) nodes. This indicate that the top four nouns co-appear with many other nouns which in turn frequently co-appear with each other while the majority of the nouns appear in very few verses. This is also established from the distribution of ACS shown in Fig. 5b. It is worth noting that about 93 percent of the nodes core score is less than the one standard deviation (1σ) above the mean with average degree and strength as 6.62 and 9.18, respectively. The average degree, 10.3 and the average strength 18.5 of the overall network are mainly attributed due to the nodes with top 7 percent ACS.

We further compute the correlation between the ACS and other methods. The results show that ACS and AR has highest correlation (0.82), followed by ACS and s-core with 0.77 correlation score and ACS and Kshell with 0.75 correlation score. This further indicates that the high degree nodes in KJB network tend to have higher strength and share a densely and strongly connected neighborhood.

TABLE 3
Nouns of KJB Network with Top-15 ACS

Proper Nouns	ACS	Proper Nouns	ACS
Israel	1	Gibeon	0.303
Judah	0.693	Dan	0.298
David	0.538	Lot	0.296
Jerusalem	0.499	Manasseh	0.29
Ephraim	0.357	Jordan	0.288
Egypt	0.353	Galilee	0.287
Benjamin	0.312	Moses	0.285
Gilead	0.31		

6.3 Social Networks

An online social network constructed from the dataset of an online Facebook community for students at University of California, Irvine [51]. The original network with self-loop was converted to a simple undirected network. If i, j is an directed edge in the original network with weight w_{ij}, then an undirected edge i, j is constructed in the transformed network with weight $w_{i,j}$. The multiple edges were then replaced with a single edge whose weight is the sum of the multiple edge weights. The undirected network indicates at least one message was exchanged across every tie. The summary statistics of the data sets are shown in the Table 1. Understanding the core nodes can reveal better insight about the relationship the students at University of California, Irvine share. Due to privacy reasons the identity of the users are kept undisclosed.

Fig. 5b shows the distribution of ACS on OSN. From the figure it is apparent that there exist a clear core structure in the network as there exist a sharp fall in the core score distribution.

In Fig. S.3, available online, we further analyze the connectivity patterns of the top ACS nodes in OSN. Fig. S.3a, available online, shows the neighborhood of the top ACS node. As evident from the figure, the node also plays the role of a hub. Moreover, about half of the neighbors of the top ACS node are not connected to any other common neighbor. This is a likely situation when the network is still in the growing phase or the pendent nodes are rather newly joined nodes. Fig. S.3b, available online, shows the neighborhood of the top 30 ACS nodes. The figure clearly shows a layered architecture where the core nodes are the densely and strongly connected nodes. Fig. S.3c, available online, shows the interconnection amongst the top 30 ACS nodes. It is evident from the figure that either the nodes are connected with high weight edges or share many edges. The figure also shows that the core structure can be of irregular shape and allow certain dissimilarity amongst nodes. In Fig. S.3d, available online, we compare the proposed core with the Kshell core. The substructure shown in the figure is the densest core identified by KShell. As shown in the figure, our method further shows a wide variation in the ACS of the KShell core. The figure clearly shows the high scoring nodes share higher strength and they are also connected to high scoring neighbors. This shows a continuous rank like ACS can provide better insight about the structures when compared to the KShell. From Fig. S.3a, available online, we also find there are many nodes who share similar core score and are classified as role equivalent.

Fig. 6b shows the correlation between ACS and various other core detection methods. As shown by the bargraphs, none of the methods are correlated to ACS. Thus we claim that ACS uncovers a very different set of core structure and node ranking as compared to the existing approaches.

6.4 Transportation Network

We next analyze a transportation network in order to determine how robust the network or a part of network is to random failure. We construct a network from the London Underground (Tube) transportation data, collected from the website for the London Underground (http://www.tfl.gov.

TABLE 4
Stations of TFL with Top 10 ACS

Station	Score	Station	Score
Leicester Square	0.702	Baker Street	0.752
Green Park	0.715	Charing Cross	0.755
King's Cross St. Pancras	0.728	Westminster	0.80
Earl's Court	0.736	Waterloo	0.971
Picadilly Circus	0.751	Bank	1

uk). In the assembled Tube network (TFL) a node represent a station, and an edge indicate that the two incident stations are adjacent on at least one line. The edge weight is calculated as the product of the number of lines directly and contiguously connecting the two stations and the inverse of the distance between the two stations. The weight function is chosen to prioritize the stations that are located in close proximity and have higher number of alternative lines connecting them. Table 1 shows the basic global properties of the network.

Table 4, show the top 10 ranking stations and their corresponding ACS. It is important to note that all the top 10 ranking stations are located in central London and participate in at least one triad. Fig. 5b shows the distribution of the core score. As shown in the figure, the ACS distribution observes a sharp drop. This indicate that a clear core structure exist in the network. Analyzing the distribution we also find that about 50 stations form the densest core while the rest 257 stations have very low core score and do not participate in core.

Additionally, we note that the degree or strength based ordering also make it possible to distinguish between core and non-core nodes. However, the ranking amongst the core nodes is much different when compared to that of the proposed method. For example, in degree based ordering King's Cross St. Pancras and Baker's Street are the top scoring stations, while following the ACS, Bank, Waterloo, Westminster score higher. This is because Bank, Waterloo, Westminster participates in triads and their adjacent stations also have multiple alternatives. King's Cross St. Pancras is more of a hub node. Fig. S.4, available online, shows how the other core nodes are positioned. It is apparent from the figure that the core analysis reveal the backbone stations of the London tube network.

6.5 Scientific Collaboration Networks

The structural analysis of scientific collaboration network, lead to better insight about the strength of the collaborations or conflicts. For the present study, we consider two author collaboration networks, where the nodes represent scholars and two scholars are linked with an edge, if they co-authored a research article. One of the networks represents the collaboration between the scholars who study computational geometry (Geom) is produced from the BibTeX bibliography [52] obtained from the Computational Geometry Database geombib, version February 2002. The weight of an edge in Geom is the number of common works between the corresponding authors. The other network represents the collaboration between the scholars who study high-energy theory in Physics and posted their pre-print work at www.arxiv.orgduring 1995-1999 [9], [53]. The edge weight in the second network is based on the number of papers coauthored by the pairs of scientists, and the number of other scientists with whom they coauthored those papers. Table 1 shows the basic global statistics of the networks, Geom and HepTh. Here we use both the continuous (β, γ) cores and the ACS ranking for the analysis.

In Fig. 7 we show the distribution of the β, γ cores. As shown in the Fig. 7, we partition the β, γ space into five zones based on the mean and one standard deviation (1σ) point above mean of the β and γ scores respectively. We further analyze the properties of the cores corresponding to various zones. In zone C both the β and γ scores are higher than the 1σ points and all the nodes lying in zone C are expected to participate in core. The zone E contains the nodes having high β and low γ scores - representing large

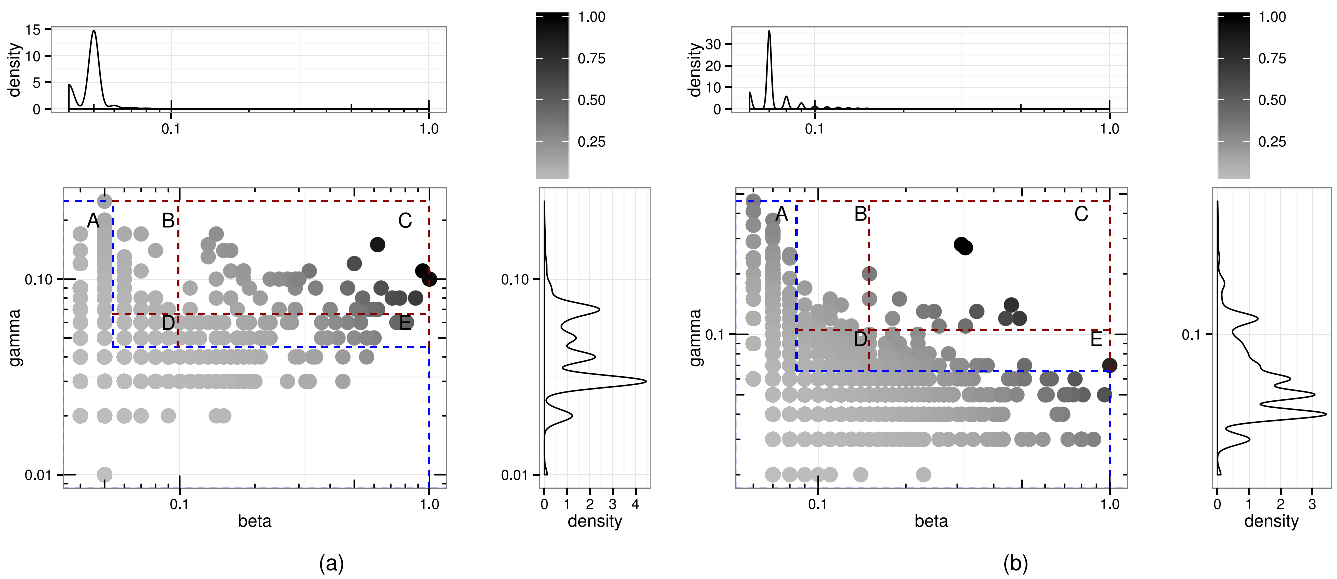

Fig. 7. The β, γ score distribution on (a) Geom and (b) HepTh network nodes. The density plots on the top and right axis shows the distribution of the β and γ scores, respectively. The colors of the nodes indicate their ACS value with the darkest colored node having an ACS of 1 and the lightest having ACS 0.

but comparatively low weight density neighborhood. The zone B contains the nodes with rather small but high weight density neighborhood indicated by low β and high γ scores. The zone A typically can be categorized as non-core having low β and γ scores. Zone D represents the intermediary structures between the cores and non-cores. The figure also indicate that the (β, γ) core are almost identical with the core obtained from ACS rank.

The top five ACS and (β, γ) ranking core nodes of zone C in Geom network are Leonidas J. Guibas of Stanford University, Pankaj K. Agarwal of Duke University, Micha Sharir of Tel Aviv University, Jack Scott Snoeyink of Stanford University, and Mark H. Overmar of Utrecht University. Guibas is also the top ranking node in terms of degree. It is interesting to note that though Sharir has coauthored the maximum number of articles, his coauthors are less strongly and less densely connected with each other as compared to that of Guibas and Agarwal. Patkos, due to the same reason, even after contributing more articles than Agarwal, ranks lower.

In the HepTh network, the two most important nodes in zone C are C.N. Pope of Center for Theoretical Physics, Texas A&M University, and H. Lu of Department of Physics and Astronomy, University of Pennsylvania. As revealed by the study, the two authors also share maximum degree and strength. When verified with www.arxiv.orgwe found that they coauthored about 69 manuscripts, indicating a strong collaboration. Also, they share a neighborhood with the maximum weight density. However, J. Ambjørnb of The Niels Bohr Institute, Denmark in spite of having maximum degree 50 and contributing about about 47 papers, scores lower as his neighbors are less strongly connected. A few important nodes from zone B are E. Elizalde of Spain, S. Nojiri of Japan, M.N. Chernodub of Russia, D.V. Nanopoulos, of Texas A&M and D. Waldram of CERN, Switzerland have rather small but strongly connected neighborhood. Two important nodes from zone E are S. Ferrara of CERN, Switzerland and P. Fre of Italy share a comparatively large neighborhood with low weight density. The study of Geom and HepTh dataset reveals an interesting collaboration pattern between several organizations across continents, showing that it is not always true that the top most contributing scientist form the most stable collaboration.

The distributions of β and γ scores as shown in the Fig. 7 also shows that the majority of the nodes are part of non-core structure while a small fraction forms the cores. In Fig. 6 we show the correlation between various methods of core decomposition. As shown by the correlation values, none of the existing methods are strongly correlated with the proposed core, indicating the proposed core reveal a unique ranking or core structure.

7 Summary And Conclusions

Meso-scale analysis of network structure has revealed many insights about the relationships between of entities of complex systems. Core decomposition is an important meso-scale analysis problem which can be applied to a broad range of domains to uncover the relationship between the network structure and node functions. The existing core decomposition algorithms based on the node properties like degree, strength, suffer from many limitations. These are typically sensitive to random noise, detect self-similar structures as cores, ignore tie-strengths, gives limited information about the connectivity of the neighborhood of a node on weighted networks. Other approaches based on the properties of a subgraph, like quasi-clique decomposition are computationally hard. The approximation algorithms are extremely sensitive to the order in which the vertices are processed and produce very different graph partitions for different vertex ordering. In addition to this, usage of various heuristics, may lead to different partitions of similar quality where there is no reason to prefer one above another. Thus the existing algorithms fail to satisfactorily reveal the stable cores of a complex heterogeneous weighted network that serve as the key organizing component of the network structure.

In the current work, we formulate a novel definition of stable core of weighted complex network, borrowing the game theoretic concept of voting mechanism. The game-theoretic methods consider the entities as rational and emphasizes on the self-organization of the entities formed based on certain purposes, not for optimizing any global objectives. We propose a game framework based on the idea that nodes vote and select a core structure by majority. In the proposed game, every node i assigns a γ score to its bounded neighbors based on their contribution in forming a strongly connected substructure. Finally an arbiter computes the co-appearance β score from the assigned γ scores. This paper proposes a set of flexible post processing approaches to cluster the nodes into cores based on their (β, γ) scores or rank the nodes based on an alternative core score derived from the (β, γ) scores. The proposed core reveals the nodes that are strongly and densely connected with high edge weights. While the alternative core score ranks the nodes based on their potential in forming strongly and densely connected structures with high edge weights. The approach favors the group of nodes that are most likely to occur together in a group. Thus the proposed core is not affected by random noise. In addition as every node vote for their preferences, the proposed method is also not affected by the vertex ordering.

An interesting property of the core is that it follows hierarchical structure. Nodes forming a dense core are always part of a rather sparse or lower core. This property helps to better identify the network organization and gives a very useful multi-scale view of the network.

To study the correctness of the developed core analysis method, we propose a synthetic benchmark with imposed core structure. This is another significant contribution of the paper that helps to study and compare the performances of core decomposition algorithms under various parameter configurations. Some of the important parameters that controls the core and non-core structures are - mixing parameter, the average degree ratio of non-core and core, the correlation between degree and strength and that between strength and clustering coefficient. Varying these parameters help in a critical analysis of the algorithms. We carried out the experiments on several benchmark instances. The results show the proposed method outperforms the state-of-the-art in almost all benchmark instances. We further show that our method can clearly discriminate between a null model with no core structure and a network with core structures from the distribution of the core values.

To evaluate the stability of the core structures, we perturb the synthetic benchmark data sets with random noise. The noise added over the weight of the edges, initially equal to w_{ij}, is equally distributed between $[-\alpha\, w_{ij}, \alpha\, w_{ij}]$, where $0 < \alpha < 1$. We compared the cores of the perturbed networks with that of the original network by computing NMI, false positive and false negative. The average value of NMI shows that the core structure remains nearly unchanged for $\alpha \leq 0.1$ and decreases only up to 25 percent for $\alpha \leq 0.5$. When compared with the existing methods, the proposed method results in maximum NMI, minimum false positive, minimum false negative with minimum variation in all three parameters. This finding strongly supports the fact that the proposed core model is more stable and favorable over the existing core models, particularly in noisy environment.

We further apply the proposed core decomposition approach for the structural analysis of the popular weighted network models WE, WPS and WSBM and weighted real life networks including character co-appearance networks of Lesmiserable, King James Bible, transportation network of London Underground systems, scientific collaboration networks on computational geometry and high energy theory, Physics, and Facebook online social network of students of University of California. The experimental results show that the proposed core analysis technique reveal important nodes that play critical roles in the organization of the networks and their hierarchical structure. The study, thus help to gain better insight about the underlying systems. The study also reveal that the core structures of the network models are different from the real networks. For example, though the WPS network model follows a power-law degree and weight distribution, the corresponding ACS tend to follow log-normal distribution while in the real networks ACS follow power-law distribution. This is possibly because of the heterogeneous network organization and sparsity of the real networks as compared to the more homogeneous organization of the network models.

In future, it would be interesting to study the centrality of the nodes identified by the dense cores. Another interesting line of future research would be to study the evolution of core structures on dynamic networks.

ACKNOWLEDGMENTS

The authors express their sincere gratitude to the anonymous reviewers for their valuable suggestions that helped to improve the quality of the manuscript. The authors are also thankful to Dr. Antonios Garas for sharing the source code of k-shell decomposition method.

REFERENCES

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Phys. Rep.*, vol. 424, no. 45, pp. 175–308, 2006.

[2] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS Biol.*, vol. 6, no. 7, p. 159, 2008.

[3] M. Ekman, J. Derrfuss, M. Tittgemeyer, and C. J. Fiebach, "Predicting errors from reconfiguration patterns in human brain networks," *Proc. Nat. Acad. Sci.*, vol. 109, no. 41, pp. 16 714–16 719, 2012.

[4] R. Sharan, I. Ulitsky, and R. Shamir, "Network-based prediction of protein function," *Molecular Syst. Biol.*, vol. 3, no. 1, 2007, Doi: 10.1038/msb4100129.

[5] G. D. Bader and C. W. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," *BMC Bioinformat.*, vol. 4, no. 1, p. 2, 2003.

[6] J.-H. Lin, Q. Guo, W.-Z. Dong, L.-Y. Tang, and J.-G. Liu, "Identifying the node spreading influence with largest k-core values," *Phys. Lett. A*, vol. 378, no. 45, pp. 3279–3284, 2014.

[7] B. Corominas-Murtra, B. Fuchs, and S. Thurner, "Detection of the elite structure in a virtual multiplex social system by means of a generalised k-core," *PloS One*, vol. 9, no. 12, p. e112606, 2014.

[8] C. Giatsidis, D. M. Thilikos, and M. Vazirgiannis, "Evaluating cooperation in communities with the k-core structure," in *Proc. Int. Conf. Adv. Soc. Netw. Anal. Mining*, 2011, pp. 87–93.

[9] M. Newman, "The structure of scientific collaboration networks," *Proc. Nat. Acad. Sci.*, vol. 98, no. 2, pp. 404–409, 2001.

[10] J. Moody, "The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999," *Am. Sociol. Rev.*, vol. 69, no. 2, pp. 213–238, 2004.

[11] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli, "Debtrank: Too central to fail? Financial networks, the fed and systemic risk," *Sci. Rep.*, Nature publishing group, vol. 2, 2012, Doi: 10.1038/srep00541.

[12] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, "k-core decomposition of internet graphs: Hierarchies, self-similarity and measurement biases," *arXiv preprint cs/0511007*, 2005.

[13] D. R. Wuellner, S. Roy, and R. M. DSouza, "Resilience and rewiring of the passenger airline networks in the united states," *Phys. Rev. E*, vol. 82, no. 5, p. 056101, 2010.

[14] Z. Xiangyu, L. Feng, Y. Rui, Z. XueMin, M. Shengwei, Z. Zhen'an, and L. Xiaomeng, "Identification of key transmission lines in power grid using modified k-core decomposition," in *Proc. 3rd Int. Conf. Elect. Power Energy Conversion Syst.*, 2013, pp. 1–6.

[15] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, "k-core decomposition: A tool for the visualization of large scale networks," *arXiv preprint cs/0504107*, 2005.

[16] S. B. Seidman, "Network structure and minimum degree," *Soc. Netw.*, vol. 5, no. 3, pp. 269–287, 1983.

[17] M. Eidsaa and E. Almaas, "s-core network decomposition: A generalization of k-core analysis to weighted networks," *Phys. Rev. E*, vol. 88, p. 062819, Dec. 2013.

[18] S. Dorogovtsev, A. Goltsev, and J. Mendes, "k-core organization of complex networks," *Phys. Rev. Lett.*, vol. 96, pp. 040 601–040 604, Feb 2006.

[19] A. Adiga and A. Vullikanti, "How robust is the core of a network?" in *Proc. Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, 2013, vol. 8188, pp. 541–556.

[20] A. Garas, F. Schweitzer, and S. Havlin, "A k-shell decomposition method for weighted networks," *New J. Phys.*, vol. 14, no. 8, p. 083030, 2012.

[21] H. Alvari, S. Hashemi, and A. Hamzeh, "Discovering overlapping communities in social networks: A novel game-theoretic approach," *AI Commun.*, vol. 26, no. 2, pp. 161–177, 2013.

[22] S. H. Yook, H. Jeong, A.-L. Barabási, and Y. Tu, "Weighted evolving networks," *Phys. Rev. Lett.*, vol. 86, pp. 5835–5838, 2001.

[23] A. Barrat, M. Barthélemy, and A. Vespignani, "Weighted evolving networks: coupling topology and weight dynamics," *Phys. Rev. Lett.*, vol. 92, no. 22, p. 228701, 2004.

[24] C. Aicher, A. Z. Jacobs, and A. Clauset. (2013). Adapting the stochastic block model to edge-weighted networks. *CoRR*, vol. abs/1305.5782, 2013 [Online]. Available: http://arxiv.org/abs/1305.5782

[25] R. G. Michael and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1979.

[26] D. Aloise, S. Cafieri, G. Caporossi, P. Hansen, S. Perron, and L. Liberti, "Column generation algorithms for exact modularity maximization in networks," *Phys. Rev. E*, vol. 82, pp. 046 112–046 120, 2010.

[27] W. Chen, A. Dress, and W. Yu, "Community structures of networks," *Math. Comput. Sci.*, vol. 1, no. 3, pp. 441–457, 2008.

[28] J. Pattillo, N. Youssef, and S. Butenko, "Clique relaxation models in social network analysis," in *Handbook Optimization in Complex Netw.*, New York. NY, USA: Springer, pp. 143–162, 2012.

[29] V. Batagelj and M. Zaversnik. (2003). An o(m) algorithm for cores decomposition of networks. *CoRR* [Online]. *cs.DS/0310049*. Available: http://arxiv.org/abs/cs.DS/0310049

[30] R.-H. Li, J. Yu, and R. Mao, "Efficient core maintenance in large dynamic graphs," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 10, pp. 2453–2465, Oct. 2014.

[31] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, "Lethality and centrality in protein networks," *Nature*, vol. 411, no. 6833, pp. 41–42, 2001.

[32] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, Jan. 2002.

[33] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, "Large scale networks fingerprinting and visualization using the k-core decomposition," in *Proc. Adv. Neural Inf. Process. Syst.*, 2005, pp. 41–50.

[34] G. D. Bader and C. W. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," *BMC Bioinformat.*, vol. 4, no. 1, p. 2, 2003.

[35] S. Wachi, K. Yoneda, and R. Wu, "Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues," *Bioinformatics*, vol. 21, no. 23, pp. 4205–4208, 2005.

[36] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, "Identification of influential spreaders in complex networks," *Nature Phys.*, vol. 6, no. 11, pp. 888–893, 2010.

[37] D. Garcia, P. Mavrodiev, and F. Schweitzer, "Social resilience in online communities: The autopsy of friendster," in *Proc. 1st ACM Conf. Online Soc. Netw.*, 2013, pp. 39–50.

[38] J. Pattillo, A. Veremyev, S. Butenko, and V. Boginski, "On the maximum quasi-clique problem," *Discrete Appl. Math.*, vol. 161, no. 12, pp. 244–257, 2013.

[39] T. Uno, "An efficient algorithm for enumerating pseudo cliques," in *Proc. 18th Int. Conf. Algorithms Comput.*, 2007, vol. 4835, pp. 402–414.

[40] J. Abello, M. Resende, and S. Sudarsky, "Massive quasi-clique detection," in *LATIN 2002: Theoretical Informatics*, S. Rajsbaum, Ed. Berlin, Germany: Springer, 2002, vol. 2286, pp. 598–612.

[41] F. Mahdavi Pajouh, Z. Miao, and B. Balasundaram, "A branch-and-bound approach for maximum quasi-cliques," *Ann. Oper. Res.*, pp. 1–17, 2012.

[42] C. Tsourakakis, F. Bonchi, A. Gionis, F. Gullo, and M. Tsiarli, "Denser than the densest subgraph: Extracting optimal quasi-cliques with quality guarantees," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 104–112.

[43] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *J. ACM*, vol. 46, no. 5, pp. 604–632, 1999.

[44] D. Wang, C. Song, and A.-L. Barabási, "Quantifying long-term scientific impact," *Science*, vol. 342, no. 6154, pp. 127–132, 2013.

[45] H. Liao, R. Xiao, G. Cimini, and M. Medo, "Network-driven reputation in online scientific communities," *PloS One*, vol. 9, no. 12, p. e112022, 2014.

[46] D. Gfeller, J.-C. Chappelier, and P. De Los Rios, "Finding instabilities in the community structure of complex networks," *Phys. Rev. E*, vol. 72, pp. 056 135–056 140, Nov. 2005.

[47] L. Ana and A. Jain, "Robust data clustering," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, Jun. 2003, vol. 2, pp. II–128–II–133.

[48] D. E. Knuth, *The Stanford GraphBase: A Platform for Combinatorial Computing*. Reading, MA, USA: Addison-Wesley, 1993, vol. 37.

[49] J. Kunegis, "Konect: The koblenz network collection," in *Proc. 22nd Int. Conf. World Wide Web Companion*, 2013, pp. 1343–1350.

[50] (2014, Oct.). King james network dataset – KONECT[Online]. Available: http://konect.uni-koblenz.de/networks/moreno_names

[51] T. Opsahl and P. Panzarasa, "Clustering in weighted networks," *Soc. Netw.*, vol. 31, no. 2, pp. 155–163, 2009.

[52] N. Beebe. (2002). Nelson h.f. beebe's bibliographies page [Online]. Available: http://www.math.utah.edu/ beebe/bibliographies.html

[53] M. Newman, "Scientific collaboration networks ii. shortest paths, weighted networks, and centrality," *Phys. Rev. E*, vol. 64, pp. 016 132–016 138, Jun. 2001.

Srinka Basu received the BTech degree in computer science from Kalyani University, India, in 2004. She received the ME degree in computer science from Jadavpur University, Kolkata, India, in 2014. She worked in the organizations like SAPLabs, India, IBM India Pvt. Ltd for about seven years as a consultant. She is working as a research fellow in the Department of Computer Science and Engineering at the Jadavpur University, Kolkata, India. Her research interests include social network analysis, game theory, machine learning, combinatorial optimization, and graph theory.

Ujjwal Maulik received the bachelor's degree in physics and computer science in 1986 and 1989, respectively. Subsequently, he received the master's and PhD degrees in computer science in 1992 and 1997, respectively. He is a professor in the Department of Computer Science and Engineering, Jadavpur University, Kolkata, India, since 2004. He chaired the Department of Computer Science and Technology Kalyani Govt. Engg. College, Kalyani, India, during 1996 to 1999. He worked at the Los Alamos National Laboratory, Los Alamos, New Mexico, in 1997, University of New South Wales, Sydney, Australia, in 1999, University of Texas at Arlington in 2001, University of Maryland Baltimore County in 2004, Fraunhofer Institute AiS, St. Augustin, Germany, in 2005, Tsinghua University, China, in 2007, University of Rome, Italy, in 2008, University of Heidelberg, Germany, in 2009, German Cancer Research Center (DKFZ) in 2010, 2011, and 2012, Grenoble INP, France, in 2010 and 2013, and ICM, Univ. of Warsaw, Poland, in 2013, International Center of Theoretical Physics (ICTP), Trieste, Italy, in 2014, University of Padoa in 2014, Grenoble INP, France, in 2014 and Corvisus University Budapest, Hungary, 2015. He has also visited many Institutes/Universities around the world for invited lectures and collaborative research. He has been invited to supervise doctoral students at a well known university in France. He is a coauthor of seven books and more than 250 research publications. He received the Govt. of India BOYSCAST fellowship in 2001, Alexander Von Humboldt Fellowship for experienced researchers in 2010, 2011, and 2012 and a senior associate of ICTP, Italy in 2012. He coordinators five Erasmus Mundus Mobility with Asia (EMMA) programs (European-Asian mobility program). He has been the program chair, tutorial chair, and member of the program committee of many international conferences and workshops. He is the associate editor of the *IEEE Transactions on Fuzzy Systems* and *Information Sciences* and also on the editorial board of many journals including *Protein & Peptide Letters*. Moreover, he has also served as guest coeditor of special issues of journals including the *IEEE Transactions on Evolutionary Computation*. He is the founder member of IEEE Computational Intelligence Society (CIS) Chapter, Kolkata section and has worked as a secretary and treasurer in 2011, as vice chair in 2012, and as the chair in 2013 as well as 2014. His research interests include computational intelligence, bioinformatics, combinatorial optimization, pattern recognition, and data mining. He is a fellow of the Indian National Academy of Engineering, West Bengal Association of Science and Technology, Institution of Engineering and Telecommunication Engineers, Institution of Engineers, and a senior member of the IEEE.