



HAL
open science

Fixed Points of Graph Peeling

James Abello, François Queyroi

► **To cite this version:**

James Abello, François Queyroi. Fixed Points of Graph Peeling. 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Aug 2013, Niagara Falls, Canada. hal-00832807

HAL Id: hal-00832807

<https://hal.science/hal-00832807v1>

Submitted on 11 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fixed Points of Graph Peeling

James Abello

DIMACS Center, Rutgers University

Piscataway, NJ, USA

Email: abello@dimacs.rutgers.edu

François Queyroi

LaBRI – University of Bordeaux – CNRS

Bordeaux, France

Email: francois.queyroi@labri.fr

Abstract—Degree peeling is used to study complex networks. It corresponds to a decomposition of the graph into vertex groups of increasing minimum degree. However, the peeling value of a vertex is non-local in this context since it relies on the connections the vertex has to groups above it. We explore a different way to decompose a network into edge layers such that the *local* peeling value of the vertices on each layer does not depend on their non-local connections with the other layers. This corresponds to the decomposition of a graph into subgraphs that are invariant with respect to degree peeling, i.e. they are fixed points. We introduce in this context a method to partition the edges of a graph into fixed points of degree peeling, called the iterative-edge-core decomposition. Information from this decomposition is used to formulate a notion of vertex diversity based on Shannon’s entropy. We illustrate the usefulness of this decomposition in social network analysis. Our method can be used for community detection and graph visualization.

I. INTRODUCTION

The peeling value of a node v in a network is the largest k such that v belongs to a maximal induced subgraph with minimum degree k [1]. In social networks, maximal induced subgraphs with peeling value at least k may be interpreted as some form of equilibrium for “a model of user engagement”. In this scenario, “each player incurs a cost to remain engaged but derives a benefit proportional to the number of engaged neighbors”[2].

The peeling value was studied for random graphs[3] generated using the Erdős-Rényi model[4]. The maximum peeling value of a graph (also called *degeneracy*) relates to other graph theoretical measures such as the coloring number[5] and arboricity. In [6], a peeling ordering of the vertices is used to improve the running time of an algorithm for the maximal cliques problem.

Degree peeling or concepts related to it are useful in network analysis. It has been used to evaluate the relevance of communities in co-authorship networks[7]. The authors proposed a reformulation of peeling that takes into account edge weights. Some graph decompositions based on degree peeling have been used in [8] and [9] as an aid to provide *layered* visualizations of graphs. Some aspects of internet topology[10] have been addressed also in this context.

One of the interesting aspects of degree peeling is the unravelling of a network hierarchy. This hierarchy is simply obtained by partitioning the vertices of the network into groups according to their peeling value (in increasing order). The group with highest peeling value is called the core of the graph. The unique group that a vertex belongs to depends not only on the number of connections it has to vertices in its group but also on its connections to vertices in upper groups.

Contribution: In this work we exploit the inherent locality of vertex peeling to efficiently obtain not only a partition of the vertex set but more importantly a partition of the edge set of any network. The algorithm’s complexity is $\mathcal{O}(k|E|)$ where k is the maximum peeling value and $|E|$ is the number of edges. The obtained edge partition, called here the *iterative edge core decomposition*, provides simultaneously distant and close readings of a network. It can be used to examine a network at different levels of granularity without losing sight of the underlying vertex partition determined by the peel values. Each subset of edges, in the iterative edge core decomposition, defines a subgraph all of whose vertices have “local peeling value” = minimum subgraph degree. Equivalently, these subgraphs are *fixed points of degree peeling* (see Figure 2 for an example). We call each of them an *edge layer*. Since a vertex can be shared among different layers we use this information to record a vertex peeling profile. This profile is an indicator of the vertex role in the network. Its Shannon’s entropy measures the degree of involvement in the peeling community structure. We exemplify our findings on a sample of social networks.

The rest of the paper is organized as follows. Notational conventions and the basic concepts used are presented in Section II. It also illustrates some of the main characteristics of graphs that are fixed points of degree peeling. In Section III, we introduce the iterative-edge-core decomposition, its main properties, and an efficient algorithm to compute it. Section IV indicates how to use the edge core decomposition to filter and analyse a network at different scales and it proposes a measure of vertex diversity based on Shannon’s entropy. Applications of the proposed edge decomposition on a sample of social networks is the main subject of Section V. We close with a discussion of possible future research directions in Section VI.

II. PEELING VALUES AND FIXED POINTS

We use the term *network* interchangeably with *graph*. We concentrate on unweighted and undirected graphs even though most of the peeling based concepts are generalizable to weighted and directed graphs[7].

In this section, we will use the co-appearance network of *Les Misérables*[11] to illustrate the different concepts used (see Figure 1). The vertices correspond to characters of the novel of Victor Hugo and an edge connects two characters if they are found together in at least one chapter.

We denote by G an undirected graph with vertex set $V(G)$ and edge set $E(G)$. A partition of $V(G)$ is called a *vertex decomposition*. Similarly a partition of $E(G)$ is called an *edge decomposition*. The degree of a vertex u in G and the minimum degree are denoted by $d_G(u)$ and $d^-(G)$ respectively. The

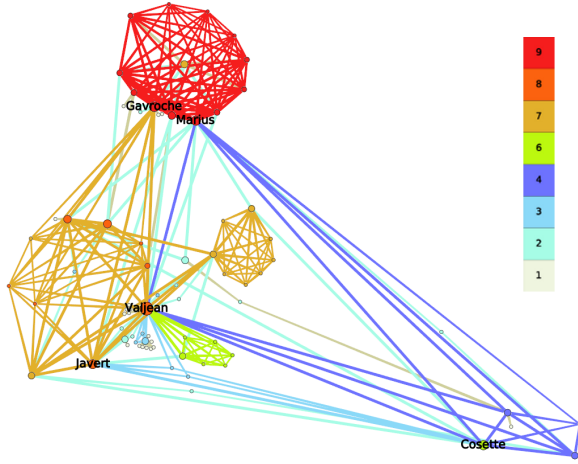


Fig. 1. The network *Les Misérables*. The peeling value is color coded. The edge coloring corresponds to our new edge decomposition (see Section III). The 5 most diverse characters (see Section IV) are explicitly labelled.

subgraph induced by a subset of vertices S is $G[S]$. For a given subset of edges $L \subseteq E$, the *layer* of G determined by L is the subgraph $G(L) = (V', L)$ where $V' = \{u \in V(G), \exists(u, v) \in L\}$.

Definition 1: (Peeling Value) The *peeling value* of a vertex $u \in V(G)$ denoted $peel_G(u)$ is the minimum value $i \in [1, d_G(u)]$ at which u is removed from G during the iterative removal of vertices of degree lower or equal to i . The *peeling value* of an edge $e \in E(G)$ denoted $peel_G(e)$ is the minimum peeling value of its endpoints (notice it is well defined in both directed and undirected cases).

In Figure 1, the peeling value of each character is mapped to the vertex color. For example, the main character “Valjean” has a peeling value of 8. The maximum peeling value of this network is 9 (red vertices).

For RAM resident graphs, the peeling value of all vertices can be computed efficiently in $\mathcal{O}(|E(G)|)$ [12]. For graphs that do not fit in RAM, an I/O efficient external-memory algorithm that computes an approximation to the peeling values has been recently proposed by [13].

The *peeling value* of G is the maximum peel value of all its vertices. The peeling value of G is also called the *degeneracy* of G [14]. For a graph of peeling value k , its vertices can be ordered in a sequence (v_1, \dots, v_n) called Erdős-Hajnal sequence[15] such that there are at most k edges going from v_i to (v_{i+1}, \dots, v_n) .

An easy but fundamental property of peel values is that they are a local manifestation of a global graph connectivity phenomenon. The following result states this precisely.

Theorem 1: (Peeling Value Locality[16]) A vertex $u \in V$ has at least $peel_G(u)$ neighbours with a peeling value at least $peel_G(u)$ and at most $peel_G(u)$ neighbours with a peeling value at least $peel_G(u) + 1$.

The authors of the previous theorem exploit these local relations between the peeling value of a vertex and the peeling values of its neighbours to compute peeling values by a distributed algorithm.

Definition 2: (Peel Decomposition and Graph Cores) The *vertex peel decomposition* of a graph G is the partition induced by the peeling values of the vertices of G . The *core* of G , $core(G)$, is the maximal subset of vertices of G whose peeling value is maximum, i.e. equal to the peeling value of G .

In Figure 1, the vertex peel decomposition is color coded by assigning the same color to all characters (vertices) with the same peeling value. This vertex partition contains 8 groups. The core of this network corresponds to the group of red characters.

Definition 3: (Local Peeling Values) Let \mathcal{P} be a partition of $V(G)$ the *local peeling value* of a vertex $u \in G$ is equal to $peel_{G[\mathcal{P}(u)]}(u)$ where $\mathcal{P}(u)$ is the set in \mathcal{P} that contains u . Similarly, if \mathcal{L} is a partition of $E(G)$, the *local peeling value* of an edge $e \in E(G)$ is equal to $peel_{G[\mathcal{L}(e)]}(e)$ where $\mathcal{L}(e)$ is the set in \mathcal{L} that contains e .

Definition 4: (Fixed Point) A graph F is a fixed point of degree peeling k if $core(F) = V(F)$ and the peeling value of F is k . Equivalently, a graph F is a *fixed point of degree peeling* if the vertex peel decomposition of F has only one class and its peeling value is equal to its minimum degree.

Note that if F is a fixed point of degree peeling, the local peeling values of elements in F do not depend on elements with higher local peeling values. Our quest is therefore to partition the edge set of a graph G into a union of fixed points of degree peeling. Among all possible edge partitions of G into fixed points of degree peeling, the one we propose is maximal (a precise definition of maximality is given in Section III). For the network *Les Misérables*, the edge coloring in Figure 1 corresponds to this decomposition. Each set of edges with the same color forms a layer of the network and this layer is a fixed point of degree peeling. The subgraph determined by brown edges corresponds for example to a fixed point of peeling value 7.

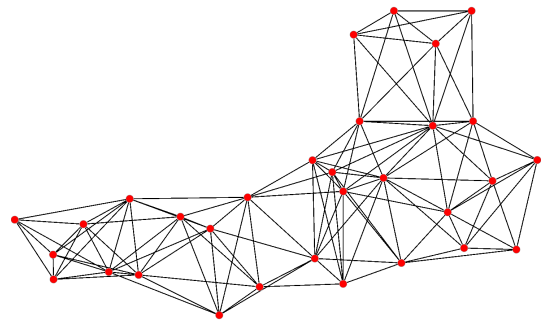


Fig. 2. An example of a random geometric subgraph in FP_5 .

We denote by FP_k the class of graphs that are fixed points of degree peeling k . They are also called *strongly k -degenerate graphs* in [17]. FP_k includes well-known classes of graphs. For example, the class FP_1 corresponds to forests (without isolated vertices), cliques of size n are in FP_{n-1} , k -regular graphs are in FP_k , and one can easily exhibit less obvious graph classes (see Figure 2). For fixed points $F \in FP_k$, the peel value locality property captured by theorem 1 can be restated as: “a vertex $u \in V(F)$ has at least k neighbours of

peeling value k ". The size of the maximum clique in $F \in FP_k$ is bounded above by $k + 1$. Bounds on the minimum and maximum number of edges of F are given in the following proposition.

Proposition 1: (E(F) Size) Let $F \in FP_k$ a fixed point of degree peeling with n vertices, we have

$$\frac{kn}{2} \leq |E(F)| \leq kn - \binom{k+1}{2} \quad (1)$$

The lower bound of Inequality (1) is the number of edges in a k -regular graph with n vertices. The upper bound is the number of edges in a edge-maximal FP_k graphs with n vertices *i.e.* graphs such that an edge can not be added between two independent vertices without increasing the maximum peeling value[17]. Graphs generated using the Barabási-Albert model[18] with a clique of size k as seed for example are in this case, edge maximal FP_k . More generally, the construction of any "edge-maximal" FP_k graph goes as follows: from a clique of size k iteratively add $(n-k)$ vertices linked to exactly k vertices. This property indicates that the average degree of an FP_k graph with n vertices is αk with $1 \leq \alpha \leq 2$. Any k -connected subgraph or connected component of an FP_k graph is a fixed point with peeling value k .

III. DECOMPOSITION INTO FIXED POINTS OF DEGREE PEELING

In this section we present two different decompositions of a graph into fixed points of degree peeling. In principle, a peeling based vertex decomposition into fixed points can be obtained by first partitioning the vertex set into groups, according to their peeling values, and then recursively applying the peel decomposition to the subgraphs induced by each set in the partition. Since the peeling value can not increase one can stop the recursion when the peeling value remains the same. In other words, the recursion will end when fixed points are reached (see an example in Figure 3(a)). This divisive decomposition is just one possible partition into fixed points. Observe however that the partition may not be maximal, in the sense that some of the obtained fixed points could be merged to obtain fixed points of higher peeling value. Notice that the same idea could be used to partition the edges. The resulting decomposition can also be non-maximal.

Among all possible vertex or edge decompositions into fixed points of degree peeling, the two we propose respect the following *maximality* property.

Definition 5: (Maximal FP decomposition) For a graph G , a vertex or edge decomposition \mathcal{P} into fixed points is said to be *maximal* iff \mathcal{P} does not contain two fixed points of the same peeling value and for any sub-collection S of \mathcal{P} whose vertex union is U , $core(G[U])$ is equal to the fixed point of maximum peeling value in S .

Notice that the maximality of an FP decomposition implies that the merging of sets in the decomposition will always result in fixed points with peeling value lesser or equal to the maximum peeling value of its components. Notice also that any vertex with local peeling value k_i has at most $(k_j - 1)$ connections to a fixed point P_j of peeling value $k_j > k_i$.

As an illustration, in Figure 3(b), we can find a larger FP_4 by adding to the red set the two vertices labelled 4 that are not in the set. Equivalently, if \mathcal{P} is a maximal edge decomposition then there is no subset of edges that can be moved to a fixed point of higher peeling value without damaging the fixed point property.

A simple and efficient approach to obtain a maximal partition into fixed points is to iteratively remove the core vertices and all its connections or alternatively remove just the edges with both end points on the core. In the first case we obtain a *vertex partition* into maximal fixed points and in the second case we obtain our desired *edge partition* into fixed points.

In both cases, we are iteratively peeling vertices or edges in the core starting with the initial graph core. We refer to both of these methods as "backward peeling". Backward vertex peeling produces what we call an "iterative vertex core decomposition" and backward edge peeling produces our desired "iterative edge core decomposition". These simple methods are formally stated below in Algorithm 1 and 2. Their complexity is $\mathcal{O}(k|E|)$ where k is the maximum peeling value of G . Their correctness follows directly from the first principle properties of the peel values stated in Section 2. Since the main focus of this paper is the edge partition into fixed points, we discuss further the properties of the iterative edge core decomposition although similar statements can be proved for the iterative vertex core decomposition. Figure 3(c) and 3(d) illustrate the differences between the iterative vertex core decomposition and the iterative edge core decomposition.

A. Iterative Vertex Core Decomposition via Backward Peeling

Algorithm 1 computes a vertex decomposition of G into fixed points of degree peeling. It relies on the fact that, for any graph G , $core(G)$ is a fixed point of degree peeling. After the removal of $core(G)$, the peeling value of the remaining vertices will directly drop if they were connected to the core. This operation can affect other vertices due to the iterative computation of peeling values. This means that in each obtained fixed point F of peeling value k , all vertices in F have a local peeling value lower or equal to their global peeling value in G .

Algorithm 1: Iterative Vertex Core decomposition of G .

Input: $G = (V, E)$
Output: $\mathcal{C} = (C_1, \dots, C_l)$, each C_i is a fixed point.

- 1 $G' \leftarrow G$;
- 2 $\mathcal{C} \leftarrow \emptyset$;
- 3 **while** $V(G') > 0$ **do**
- 4 $\mathcal{C} \leftarrow \mathcal{C} \cup \{core(G')\}$;
- 5 $G' \leftarrow G'[V(G') - core(G')]$;
- 6 **end**
- 7 **return** \mathcal{C} ;

Notice that this iterative vertex core decomposition discards the connections between the different groups in the graph. This is one of the main reasons we introduce the following iterative edge core decomposition.

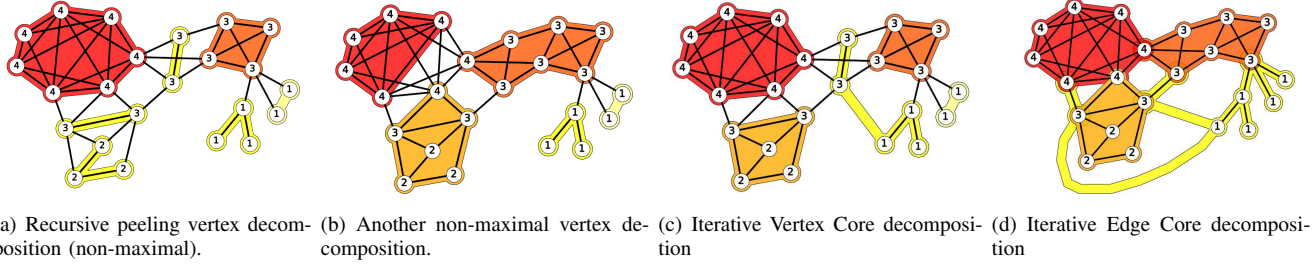


Fig. 3. Four different decompositions of a graph into fixed points of degree peeling. The induced subgraphs formed by taking all vertices in a hull is a fixed point of degree peeling. Pale yellow represents the lowest peeling value 0 and red represents the highest peeling value 4. Vertices are labeled according to their peeling value. The decomposition a) is not maximal since the subgraph induced by the union of the yellow sets has a peeling value of 2. b) is not maximal since two vertices labelled 4 have 4 connections to a FP_4 fixed point.

B. Iterative Edge Core Decomposition via Backward Peeling

The iterative edge core decomposition (see Algorithm 2) assign to each edge a value that corresponds to the peeling value of its endpoints at the first time they belong to the core. In the example given in Figure 3(d), removing the edges within the red hull leaves most of the vertices in the core isolated. Three of them have connections to the rest of the graph. The leftmost one has actually enough connections to be part of the next core (of peeling value 3) but after that it is also isolated. The idea here is that all the vertices that belong to a fixed point will not be similar, in the sense that some of them can actually be part of other fixed points.

Algorithm 2: Iterative Edge Core decomposition of G .

Input: $G = (V, E)$
Output: $\mathcal{L} = (L_1, \dots, L_p)$, each L_i is a fixed point.

- 1 $G' \leftarrow G$;
- 2 $\mathcal{L} \leftarrow \emptyset$;
- 3 **while** $E(G') > 0$ **do**
- 4 $A = \{(u, v) \in E(G'), u \in \text{core}(G') \wedge v \in \text{core}(G')\}$
- 5 $\mathcal{L} \leftarrow \mathcal{L} \cup \{A\}$;
- 6 $E(G') \leftarrow E(G') - A$;
- 7 **end**
- 8 **return** \mathcal{L} ;

The maximum number of iterations of Algorithm 2 is bounded by k . Indeed, removing the edges of the core reduces the peeling value by at least 1. The size of the resulting partition is therefore at most k . The peeling value of vertices is computed at each iteration and this operation can be done in $\mathcal{O}(|E(G)|)$. This decomposition is maximal according to Definition 5 (see Theorem 2).

Theorem 2: The edge decomposition \mathcal{L} computed by Algorithm 2 is maximal (see Property 5).

Proof: To simplify notation, we assume that each set $\mathcal{L} = (L_1, L_2, \dots, L_p)$ corresponds to the subset of vertices in the subgraph determined by each layer. First, the decomposition \mathcal{L} does not contain two fixed points with the same peeling value. Indeed, Algorithm 2 returns the fixed points formed by the core in a strictly decreasing order of their peeling value. If we assume now that \mathcal{L} is not maximal (see Definition 5). It must exist a subcollection S of \mathcal{L} whose vertex union is U and such that $\text{core}(G[U])$ is not equal to the fixed point of maximum peeling value of a fixed point in S .

We argue first that S can not contain $L_1 = \text{core}(G)$. Indeed, the set L_1 is, as core of G , the core of any induced subgraph of G that contains L_1 . Now L_1 is not in S , $L_i \not\subseteq U$. Now, since $L_1 \notin S$, the subset U is actually a subset of $(V - L_1)$. The same reasoning could be applied to show that S can not contain $L_2 = \text{core}(G[V - L_1])$, $L_3 = \text{core}(G[V - L_1 - L_2])$ etc. Finally, the subset of U must be a subset of $(V - \bigcup_{i=1}^p L_i) = \emptyset$. However there are no subset S of \mathcal{L} with an empty vertex union. Therefore, such set does not exist and \mathcal{L} is maximal. ■

IV. USES OF THE ITERATIVE EDGE CORE DECOMPOSITION

In this section, we indicate how to use the edge core decomposition to filter and analyse a network at different scales. We also propose a measure of vertex diversity based on Shannon’s entropy.

Network analysis at different scales. The iterative edge decomposition focuses on local peel values and since it is maximal each edge gets assigned to its highest possible layer. Each layer locality is captured by the fact that it is a fixed point of degree peeling. The usual peel decomposition fails to incorporate locality since the vertices of peeling value k could very well form an independent set. The peel decomposition tends to produce more layers than our iterative edge core decomposition.

Peeling based decompositions can be paraphrased using a terrain metaphor. Each local max corresponds in the network to a subset of vertices whose peeling value does not depend on the vertices with higher peeling values. The core of the network is the overall maximum. The iterative edge core decomposition follows a top-down approach. Its computation jumps from the overall maximum to subsequent levels of local “plateaux”.

Network Filtering and Community Structure. Peeling values can be used to filter out vertices with few connections. In the iterative edge core decomposition, vertices of peeling value k are present in layers with peeling value at most k . The lowest layer L_p may contain vertices of peeling value 1 but also their sparse connections between vertices from layer p or above. As an illustration, consider a network formed by two cliques of different sizes linked by an edge. This bridge edge will fall into the lowest layer of the decomposition. Some layers can be filtered out to make the community structure (if any) more apparent. By associating with each vertex the first

time that it appears in the iterative edge core decomposition, we can keep track of the proportion of recently added vertices to any layer. Sudden proportion changes between consecutive levels are an indicator of a possible community structure. Notice that the iterative edge core decomposition is more likely to detect communities than the vertex peel decomposition since edge partitioning inherently allows vertex overlaps between communities[19].

Vertex Diversity. For a given vertex u , the edges adjacent to u in G can be partitioned into different classes given by the iterative edge core decomposition. We associate with each vertex a *profile vector* containing its peeling information according to the iterative edge core decomposition.

Definition 6: For a graph G and its iterative edge core decomposition $\mathcal{L} = (L_1, \dots, L_p)$, the *profile* of a vertex $u \in V(G)$ denoted $profile(u)$ is a sequence of integers (l_1, \dots, l_p) where each l_i indicates the number of edges adjacent to u that are part of the layer L_i .

Notice that the number of times a vertex u is detected as part of a layer, in the iterative edge core decomposition, corresponds to the number of non-zero entries in the vector $profile(u)$. The sum of the entries in the profile vector of u is equal to the degree of u . Vertex profiles are used next to assess the *diversity* of a vertex.

Definition 7: (Vector Majorization[20], Shannon’s Entropy[21] and Vertex Diversity) For a vector u in R^k , let $p(u) = (p_1(u), p_2(u), \dots, p_k(u))$ denote the vector obtained by sorting the entries of u from largest to smallest. A vector v in R^k is said to be majorized by a vector u in R^k iff for $1 \leq l < k$, $\sum_{j=1}^l p_j(v) \leq \sum_{j=1}^l p_j(u)$ and $\sum_{j=1}^k p_j(v) = \sum_{j=1}^k p_j(u)$.

Let $H(profile(u))$ be the Shannon’s entropy of the profiles normalized by the degree.

$$H(profile(u)) = - \sum_{i=1}^p \frac{l_i}{d_G(u)} \log_2 \left(\frac{l_i}{d_G(u)} \right) \quad (2)$$

Notice that for two vertices u and v , if $profile(v)$ is majorized by $profile(u)$ then $H(profile(v)) \geq H(profile(u))$. Therefore, we can rank vertices using the entropy of their profiles. We call this measure vertex diversity. Namely, the *diversity of vertex u* is $H(profile(u))$.

Notice that the diversity of a vertex in this case does not solely depend on its peeling value or its highest layer in the iterative edge core decomposition. A vertex that is not part of the core of the graph can still have a bigger diversity than vertices from the core.

V. NETWORK SAMPLES

We illustrate the application of the iterative edge core decomposition on four networks with different characteristics. We use node-link diagrams generated using Tulip software[22]. In each case, the peeling values of the vertices are color coded. The color of an edge corresponds to the peeling value of its layer in the iterative edge core decomposition. Both vertices and edge values use the same color scale. The diversity of a vertex (see Eq. 2) is encoded by the size of the vertex.

TABLE I. STATISTICS FOR THE ITERATIVE EDGE CORE DECOMPOSITION OF THE MANUFACTURE NETWORK.

Layers	# Vertices	% New vertices	Avg. Clustering Coef.
16	24	100 %	0.89
11	12	100 %	1
10	11	100 %	1
8	17	88 %	0.80
5	45	31 %	0.23
2	28	4 %	0.03
1	38	0 %	0

We also provide 3D *z-ordered* visualizations of the networks by mapping into the *z*-axis the edge core decomposition numbers. This method is inspired by [9].

A. A Co-Appearance Network: Les Misérables

We start with the co-appearance network *Les Misérables* (see Section II). The network contains 77 vertices and 254 edges. The peel decomposition contains 8 groups. The core of the graph (red vertices) corresponds to the “revolutionary student club” appearing during the Paris uprising in the novel. While some of them are very important in the novel (like “Marius”), most of them are not and the reason they have the maximum peeling value is because of the size of the group. We can differentiate them by looking at their connections to vertices of lower peeling value and this is exactly what the iterative edge core provides (see Figure 1).

The second layer of the iterative edge core decomposition (brown edges) contains characters such as the Thenardier family. Their son, “Gavroche”, was part of the core but he actually has enough connections to be grouped with them at this level. Each layer seems to correspond to a community in the novel. For example, the blue edges layer determines a subgraph with 6 vertices, its density is maximum as fixed point of peeling value 4 according to Property 1. It corresponds to the group formed by Marius along with members of his family, his fiancée “Cosette” and the tutor of Cosette, “Valjean”.

The labels displayed in Figure 1 correspond to the names of the five most *diverse* characters: Valjean, Gavroche, Cosette, Marius, and Javert. Notice that Cosette which is a main character in the novel appears here even if her peeling value is relatively low when compared to the others. However, in the iterative edge core decomposition, she is well connected in the layers she belongs to.

B. A Social Interaction Network: The Manufacture Network

The example we now discuss is an intra-organizational network[23] where the vertices represent 77 employees and an edge links two employees when one of them indicates that the other provides him useful information (at least somewhat infrequently). The employees work in four different locations: Paris, Frankfurt, Warsaw and Geneva (see Figure 4).

This example illustrates the usefulness of our method for graph filtering (see Section IV). Observe a transition between the layers 8 and 5 in terms of their proportion of new vertices found and their average clustering coefficient (see Table I). The first four layers (see Figure 4(a)) separate the core of the communities induced by the locations even if they have

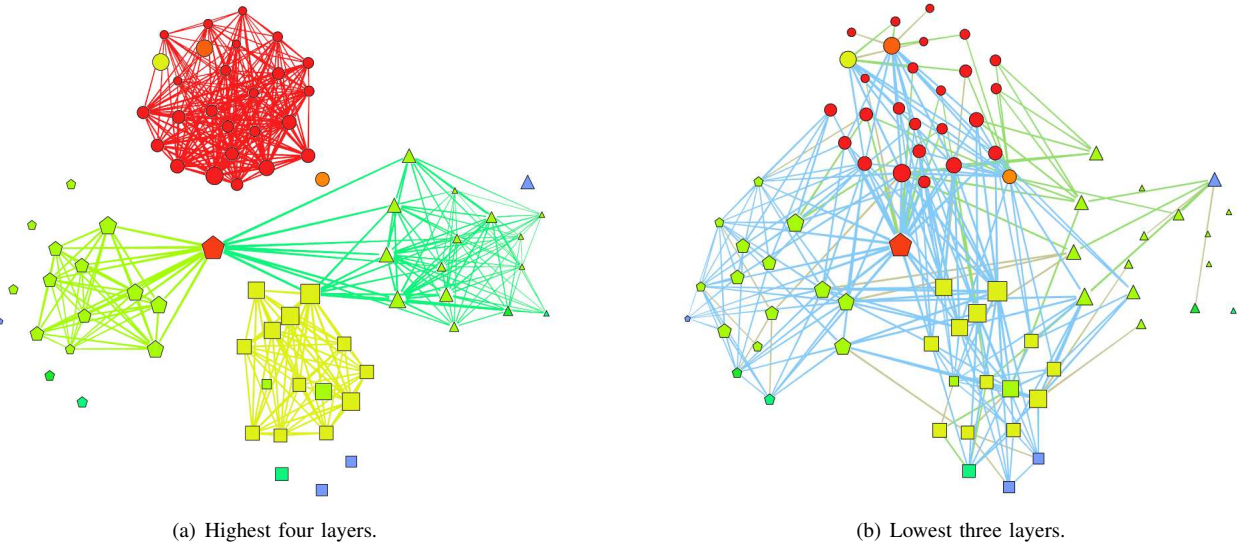


Fig. 4. Results of the iterative edge core decomposition for the Manufacture network. The graph is drawn using a force-directed layout algorithm. The positions between a) and b) are preserved. The shape of the vertices corresponds to the different locations: Frankfurt (circles), Paris (squares), Geneva (pentagons) and Warsaw (triangles). The edges of the graph are separated here according to the decomposition. The union of a) and b) gives the complete network.

different global peeling value. Two vertices in the center have also enough connections to be part of the layer of peeling value 8 that groups vertices from Warsaw. Notice that those vertices have also a high diversity. The vertices that are isolated in those layers correspond to vertices whose number of connections is too low or to employees whose connections are too split between the different locations.

The last three layers (see Figure 4(b)) also bring relevant information. The blue edges determine a fixed point of peeling value 5. This subgraph contains a substantial number of vertices from higher layers. It suggests that even without the connections between people from the same location, the graph structure still allows the diffusion of information in the company. Notice that most of the employees from Warsaw do not belong to this subgraph. This suggests that connectivity is “stronger” between people working in Frankfurt, Paris and Geneva.

C. The Political Blogosphere Network

This network represents the undirected links between political blogs before the 2004 US election[24]. The 1222 blogs are divided into two groups: liberals or conservatives (see Figure 5). The peel decomposition contains 35 clusters (the maximum peeling value in the network). The core of the network contains liberals blogs, vertices of peeling value between 32 and 34 are other liberal blogs connected to the core. A substantial change happens when looking at vertices with peeling value 31 since a large proportion of them are conservative blogs. Subsequent groups in the peel decomposition contain both liberal and conservatives blogs.

On the other hand, a quite different picture emerges using the 10 layers of the iterative edge core decomposition of this network (see Figure 5(a)). Indeed, either liberal or conservative blogs are over-represented in the layers detected (see the statistics provided in Table II). A 3D *z-ordered* visualization (see Figure 5(b)) reveals this phenomenon. Notice however, that an edge layer can contain blogs from the “opposite” side. This suggests that the local peeling values of the blogs mostly

come from blogs with the same affiliation. Observe that the average clustering coefficient of the layers decreases heavily. This can be explained by the fact that each layer contains an important proportion of vertices already detected in the previous layers.

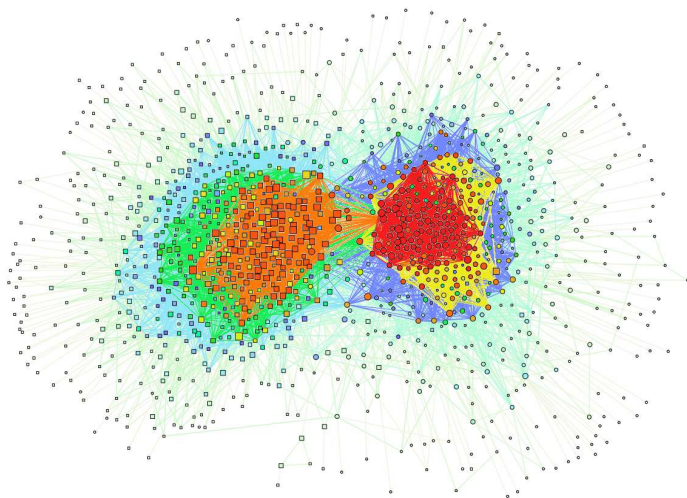
TABLE II. STATISTICS FOR THE ITERATIVE EDGE CORE DECOMPOSITION OF THE POLITICAL BLOGS NETWORK.

Layers	# Vertices	% Liberals	% New vertices	Avg. Clust. Coef.
33	66	100 %	100 %	0.72
29	95	4 %	99 %	0.5
24	120	99 %	68 %	0.26
18	186	5 %	65 %	0.14
13	206	93 %	44 %	0.09
9	315	6 %	55 %	0.05
5	383	77 %	42 %	0.03
3	382	2 %	35 %	0.02
2	373	64 %	34 %	0.02
1	706	45 %	25 %	0

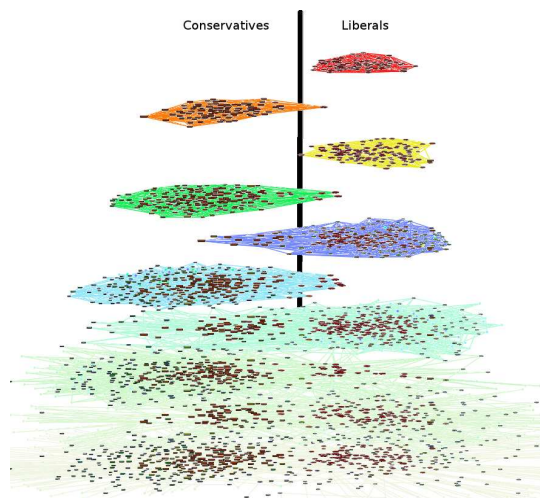
D. The Air Transport Network

The Air transport network is an undirected graph where each vertex represents the airports of a city and edges represent a direct flight from one city to the other [25]. The network contains 1490 nodes and 12353 edges. The size of the peel decomposition is equal to the maximum peeling value in this network which is 36. However, the iterative edge core decomposition provides a partition of the edges into 12 layers (see Figure 6). European cities appear in many layers although the third and sixth layers (of peeling value 17 and 8) contain a majority of Asian cities. The fact that the local peeling value is high between those airports is difficult to spot using the peel decomposition.

In this case, we notice two interesting aspects when looking at the results of the iterative edge core decomposition. The backward peeling procedure leaves a few cities isolated at

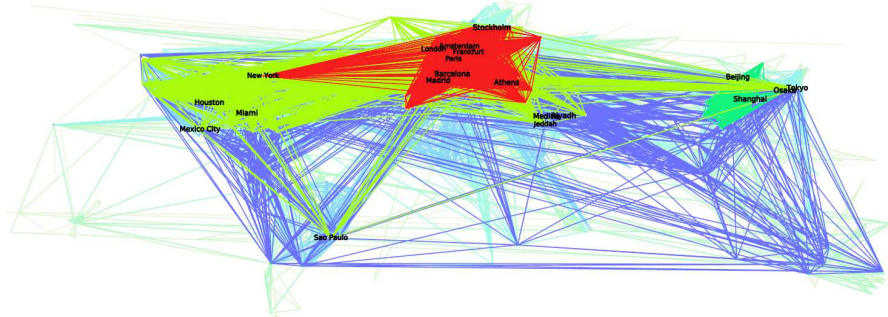


(a) Plain view obtained using a force-direct algorithm.

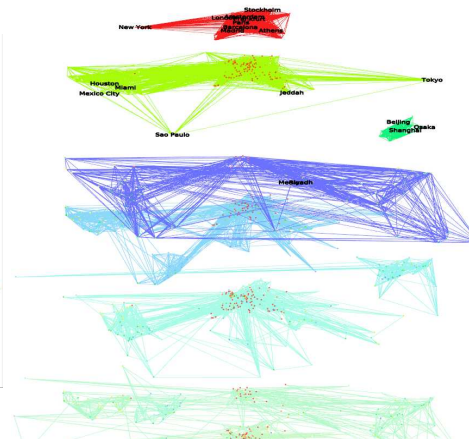


(b) 3D z -ordered view.

Fig. 5. Results of the iterative edge core decomposition for the Political Blogs network. The graph is drawn using a force-direct layout algorithm. The shape of the vertices corresponds to the politic orientation: liberal (circle) or conservative (square). The labels on the color scale indicate the peeling values of the edge layers found.



(a) Plain view obtained using geographic coordinates.



(b) z -ordered 3d view.

Fig. 6. Results of the iterative edge core decomposition for the Air Transport Network. The vertices coordinates correspond to the geographical positions to the airports. Labels indicates the names of the 20 most diverse cities according to the decomposition.

each step. As a case in point, the network remaining after the removal of all edges of local peeling value greater than 4 contains a giant connected component with 1380 airports. A second interesting aspect is observed when looking at airports diversity. The most diverse airport is Jeddah in Saudi Arabia. This may be because this city has a distinctive role being very close to the Islam’s holy city of Mecca. The next most diverse airports are: Paris, Frankfurt, London, New York City, Sao Paulo and Beijing.

We have also applied the iterative edge core decomposition to the Jazz network[26] (see Figure 7), University Facebook networks or population migration networks. We will report on these findings in a follow up paper.

VI. CONCLUSIONS

We introduced an efficient graph edge partition into fixed points of degree peeling. Each layer in the decomposition has a

unique local peel value. The presented algorithms and technics can be generalized to weighted networks. Information from the decomposition allowed us to formulate a novel notion of vertex diversity with an associated measure based on Shannon’s entropy. We illustrated graph filtering and analysis at different scales using 3D z -ordered node-link diagrams.

We are currently studying the mathematical properties of fixed points of degree peeling. For example, we want to fully characterize minimal fixed points of degree peeling. This characterization may be useful in the understanding of some fundamental graph streaming computations. Finally, our decomposition can be used as a preprocessing step for a variety of graph drawing algorithms.

ACKNOWLEDGMENT

This research is partially supported by DIMACS and mgvis.com.

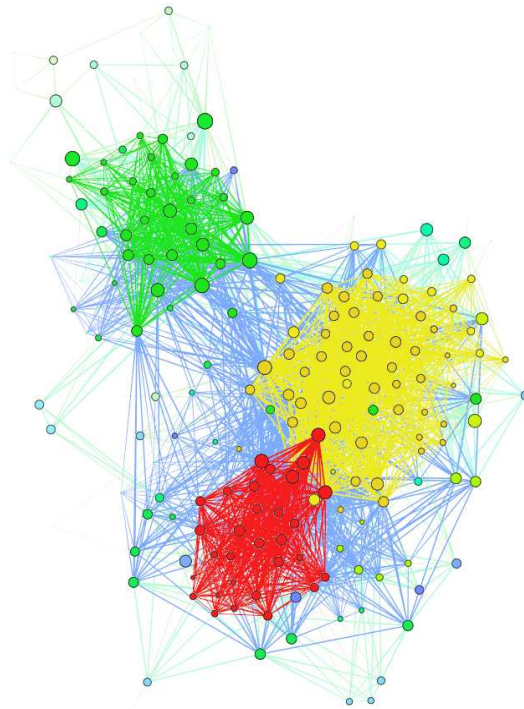


Fig. 7. Results of the iterative edge core decomposition for the Musician Jazz Network.

REFERENCES

- [1] S. Seidman, "Network structure and minimum degree," *Social networks*, vol. 5, no. 3, pp. 269–287, 1983.
- [2] K. Bhawalkar, J. Kleinberg, K. Lewi, T. Roughgarden, and A. Sharma, "Preventing unraveling in social networks: the anchored k-core problem," *Automata, Languages, and Programming*, pp. 440–451, 2012.
- [3] B. Pittel, J. Spencer, N. Wormald *et al.*, "Sudden emergence of a giant k-core in a random graph," *Journal of combinatorial theory. Series B*, vol. 67, no. 1, pp. 111–151, 1996.
- [4] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hungar. Acad. Sci.*, vol. 5, pp. 17–61, 1960.
- [5] G. Szekeres and H. S. Wilf, "An inequality for the chromatic number of a graph," *Journal of Combinatorial Theory*, vol. 4, no. 1, pp. 1–3, 1968.
- [6] D. Eppstein, M. Löffler, and D. Strash, "Listing all maximal cliques in sparse graphs in near-optimal time," *CoRR*, vol. abs/1006.5440, 2010.
- [7] C. Giatsidis, D. Thilikos, and M. Vazirgiannis, "Evaluating cooperation in communities with the k-core structure," in *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*. IEEE, 2011, pp. 87–93.
- [8] J. Alvarez-Hamelin, L. Dall Asta, A. Barrat, and A. Vespignani, "Large scale networks fingerprinting and visualization using the k-core decomposition," *Advances in neural information processing systems*, vol. 18, p. 41, 2006.
- [9] M. Baur, U. Brandes, M. Gaertler, and D. Wagner, "Drawing the as graph in 2.5 dimensions," in *Graph Drawing*. Springer, 2005, pp. 43–48.
- [10] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, "A model of internet topology using k-shell decomposition," *Proceedings of the National Academy of Sciences*, vol. 104, no. 27, pp. 11 150–11 154, 2007.
- [11] D. E. Knuth, *The Stanford GraphBase: a platform for combinatorial computing*. AcM Press, 1993.
- [12] V. Batagelj and M. Zaversnik, "An $o(m)$ algorithm for cores decomposition of networks," *arXiv preprint cs/0310049*, 2003.
- [13] M. T. Goodrich and P. Pszona, "External-memory network analysis algorithms for naturally sparse graphs," *CoRR*, vol. abs/1106.6336, 2011.
- [14] D. R. Lick and A. T. White, "k-degenerate graphs," *Canad. J. Math.*, vol. 22, pp. 1082–1096, 1970.
- [15] P. Erdős and A. Hajnal, "On chromatic number of graphs and set-systems," *Acta Mathematica Hungarica*, vol. 17, no. 1, pp. 61–99, 1966.
- [16] A. Montesor, F. D. Pellegrini, and D. Miorandi, "Distributed k-core decomposition," *CoRR*, vol. abs/1103.5320, 2011.
- [17] R. Bauer, M. Krug, and D. Wagner, "Enumerating and generating labeled k-degenerate graphs," in *7th Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, 2010, pp. 90–98.
- [18] A. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [19] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, "Link communities reveal multiscale complexity in networks," *Nature*, vol. 466, no. 7307, pp. 761–764, 2010.
- [20] B. C. Arnold, *Majorization and the Lorenz order: a brief introduction*. Springer-Verlag Berlin, 1987, vol. 43.
- [21] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.
- [22] D. Auber, D. Archambault, R. Bourqui, A. Lambert, M. Mathiaut, P. Mary, M. Delest, J. Dubois, and G. Melançon, "The Tulip 3 Framework: A Scalable Software Library for Information Visualization Applications Based on Relational Data," INRIA, Rapport de recherche RR-7860, Jan. 2012.
- [23] R. L. Cross and A. Parker, *The hidden power of social networks: Understanding how work really gets done in organizations*. Harvard Business Press, 2004.
- [24] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 us election: divided they blog," in *Proceedings of the 3rd international workshop on Link discovery*. ACM, 2005, pp. 36–43.
- [25] C. Rozenblat, G. Melançon, and P.-Y. Koenig, "Continental integration in multilevel approach of world air transportation (2000-2004)," *Networks and Spatial Economics*, 2008.
- [26] P. M. Gleiser and L. Danon, "Community structure in jazz," *Advances in complex systems*, vol. 6, no. 04, pp. 565–573, 2003.