Identifying Influential Nodes in Online Social Networks Using Principal Component Centrality

Muhammad U. Ilyas[1,2]
[1] School of Electrical Engineering & Computer Science
National University of Sciences & Technology
Islamabad, Pakistan
usman@seecs.edu.pk

Hayder Radha[2]
[2] Department of Electrical & Computer Engineering
Michigan State University
East Lansing, MI 48824, USA
{ilyasmuh, radha}@egr.msu.edu

***Abstract*—Identifying the most influential nodes in social networks is a key problem in social network analysis. However, without a strict definition of centrality the notion of what constitutes a central node in a network changes with application and the type of commodity flowing through a network. In this paper we identify *social hubs*, nodes at the center of influential neighborhoods, in massive online social networks using principal component centrality (PCC). We compare PCC with eigenvector centrality's (EVC), the *de facto* measure of node influence by virtue of their position in a network. We demonstrate PCC's performance by processing a friendship graph of $70,000$ users of Google's Orkut social networking service and a gaming graph of $143,020$ users obtained from users of Facebook's 'Fighters' Club' application.**

I. Introduction

Centrality [1], [2], [4], [9], [17] is a measure to assess the criticality of a node's position. Node centrality as a measure of a node's importance by virtue of its central location has been in common use by social scientists in the study of social networks for decades. Over the years several different meanings of centrality have emerged.

Among many centrality measures, eigenvalue centrality (EVC) is arguably the most successful tool for detecting the most influential node(s) within a social graph. Thus, EVC is a widely used centrality measure in the social sciences ([10], [19], [1], [8], [6], [7], [20], [18], [3], [2]). As we demonstrate earlier in [11], one key shortcoming of EVC is its focus on (virtually) a single influential set of nodes that tends to cluster within a single neighborhood. EVC has the tendency of identifying a set of influential nodes that are all within the same region of a graph. This shortcoming does not represent a major issue for many social science problems and Internet applications, such as PageRank, where EVC has been used extensively [13]. Meanwhile, when dealing with massive social network graphs, it is hardly the case that there is a single neighborhood of influential nodes; rather, there are usually multiple influential neighborhoods most of which are not detected or identified by EVC.

In order to identify influential neighborhoods, there is a need for a measure of centrality that identifies neighborhoods of influential nodes. One can think of a centrality plane that is overlaid over the underlying social graph. This centrality plane may contain multiple centrality score maxima, each of which is centered on an influential neighborhood. Nodes that have centrality score higher than other nodes are located under a centrality peak and are more central than any of their neighbors. We use the term *social hubs* to refer to nodes forming centrality maxima. These social hubs form the kernel of influential neighborhoods in social networks. Our focus in this paper is on identifying influential neighborhoods rather than influential nodes. To this end, we apply principal component centrality (PCC) which we introduced in [11].

We motivate the search for local centrality maxima by the following application. Suppose we wish to deploy a limited number of monitors in a social graph to spot the emergence and adoption of as many trends as possible. The spread of trends in social networks is modeled as an influence process. The degree to which a node is influential in the spread of a trend in the long term is measured by its EVC. We define a social hub in a network as a node whose centrality score forms a local maxima, i.e. its centrality score is higher than all of its neighbors'. The number of local maxima identified is used as a performance metric. The results obtained by using EVC provides the baseline for comparison. On a walk over a graph, the EVCs of nodes change gradually, i.e. a node's EVC is high in part because its neighbors EVC is also high. This means, if we were to pick nodes for placement of monitors in descending order of EVC, quite a few will end up monitoring the same well connected cluster of nodes. This will introduce redundancy in the monitoring at the cost of coverage. With a limited number of monitors available it would be more desirous to position them in different vicinities of the graph. If, on the other hand, a node is selected only if its centrality measure is significant and locally maximum we can avoid redundancy and in effect, observe a greater number of nodes with the same number of monitors. This is why our focus is on identifying influential neighborhoods rather than influential nodes.

The rest of this paper is organized as follows. Section II gives background review of EVC and PCC. Section III applies PCC to an undirected, unweighted friendship graph from Google's Orkut social networking service. Orkut currently has approximately $440,000$ active subscribers [16] and has large

This work was supported in part by NSF award IIS-0968495, NSF award CNS-0721550 and an unrestricted gift from Microsoft Research.

subscriber bases in Brazil and India. The data set available to us [14] consists of $70,000$ users connected by $2,971,776$ links. In section IV we apply PCC to a weighted, undirected gaming graph of matches between users of Facebook's 'Fighters' Club' application. This data set was originally collected by Nazir in [15]. It consists of $667,560$ recorded matches between $143,020$ users making for a graph with $526,224$ edges between users. Section V concludes the paper.

II. Background

Let $\mathbf{A}$ denote the adjacency matrix of a graph $G(V,E)$ consisting of the set of nodes $V = \{v_1, v_2, v_3, \ldots, v_N\}$ of size N and set of undirected edges E. When a link is present between two nodes v_i and v_j both $A_{i,j}$ and $A_{j,i}$ are set equal to 1 and set to 0 otherwise. Let $\Gamma(v_i)$ denote the neighborhood of v_i, the set of nodes v_i is connected to directly.

A. Eigenvector Centrality

Eigenvector centrality (EVC) is a relative score recursively defined as a function of the number and strength of connections to its neighbors and as well as those neighbors' centralities. Let $x(i)$ be the EVC score of a node v_i. Then,

$$\begin{aligned} x(i) &= \frac{1}{\lambda} \sum_{j \in \Gamma(v_i)} x(j) \\ &= \frac{1}{\lambda} \sum_{j=1}^{N} A_{i,j} x(j) \end{aligned} \tag{1}$$

Here λ is a constant. Equation 1 can be rewritten in vector form equation 2 where $\mathbf{x} = \{x(1), x(2), x(3), \ldots, x(N)\}'$ is the vector of EVC scores of all nodes.

$$\begin{aligned} \mathbf{x} &= \frac{1}{\lambda}\mathbf{A}\mathbf{x} \\ \lambda\mathbf{x} &= \mathbf{A}\mathbf{x} \end{aligned} \tag{2}$$

This is the well known eigenvector equation where this centrality takes its name from. λ is an eigenvalue and $\mathbf{x}$ is the corresponding eigenvector of matrix $\mathbf{A}$. Obviously several eigenvalue/eigenvector pairs exist for an adjacency matrix $\mathbf{A}$. The EVC of nodes are defined on the basis of the Perron eigenvalue λ_A (the Perron eigenvalue is the largest of all eigenvalues of $\mathbf{A}$ and is also called the principal eigenvalue). If λ is any other eigenvalue of $\mathbf{A}$ then $\lambda_A > |\lambda|$. The eigenvector $\mathbf{x} = \{x(1), x(2), \ldots, x(N)\}'$ corresponding to the Perron eigenvalue is the Perron eigenvector or principal eigenvector. Thus the EVC of a node v_i is the corresponding element $x(i)$ of the Perron eigenvector $\mathbf{x}$. Note that when the adjacency matrix $\mathbf{A}$ is symmetric all elements of the principal eigenvector $\mathbf{x}$ are positive.

Computing a node's EVC it takes into consideration its neighbors's EVC scores. Because of its recursive definition, EVC is suited to measure nodes' power to influence other nodes in the network both directly and indirectly through its neighbors. Connections to neighbors that are in turn well connected themselves are rated higher than connections to neighbors that are weakly connected.

B. Principal Component Centrality

While EVC assigns centrality to nodes according to the strength of the most dominant feature of the data set, PCC takes into consideration additional, subsequent features. We define the PCC of a node in a graph as the Euclidean distance/ℓ^2 norm of a node from the origin in the P-dimensional eigenspace formed by the P most significant eigenvectors. For a graph consisting of a single connected component, the N eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_N| = 0$ correspond to the normalized eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$. The eigenvector/eigenvalue pairs are indexed in order of descending magnitude of eigenvalues. When $P = 1$, PCC equals a scaled version of EVC. Unlike other measures of centrality, the parameter P in PCC can be used as a tuning parameter to adjust the number of eigenvectors included in the PCC. Let $\mathbf{X}$ denote the $N \times N$ matrix of concatenated eigenvectors $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \ldots \mathbf{x}_N]$ and let $\Lambda = [\lambda_1 \lambda_2 \ldots \lambda_N]'$ be the vector of eigenvalues. Furthermore, if $P < N$ and if matrix $\mathbf{X}$ has dimensions $N \times N$, then $\mathbf{X}_{N \times P}$ will denote the submatrix of $\mathbf{X}$ consisting of the first N rows and first P columns. Then PCC can be expressed in matrix form as:

$$\mathbf{C}_P = \sqrt{\left(\left(\mathbf{A}\mathbf{X}_{N \times P}\right) \odot \left(\mathbf{A}\mathbf{X}_{N \times P}\right)\right) \mathbf{1}_{P \times 1}} \tag{3}$$

The '$\odot$' operator is the Hadamard (or entrywise product or Schur product) operator. Equation 3 can also be written in terms of the eigenvalue and eigenvector matrices Λ and $\mathbf{X}$, of the adjacency matrix $\mathbf{A}$:

$$\mathbf{C}_P = \sqrt{\left(\mathbf{X}_{N \times P} \odot \mathbf{X}_{N \times P}\right)\left(\Lambda_{P \times 1} \odot \Lambda_{P \times 1}\right)}. \tag{4}$$

Adding successively more features/eigenvectors will have the obvious effect of increasing the sum total of node PCC scores, i.e. $\mathbf{1}_{1 \times N}\mathbf{C}_m > \mathbf{1}_{1 \times N}\mathbf{C}_n$ when $m > n$. However, it is unclear how much PCC's scores change as P is varied from 1 through N. In [5] Canright *et al.* use the phase difference between eigenvectors computed in successive iterations as a stopping criteria for their fully distributed method for computing the principal eigenvector. We use the phase angle between PCC vectors and EVC to study the effect of adding more features. We compute the phase angle $\phi(n)$ of a PCC vector using n features with the EVC vector as,

$$\phi(P) = \arccos\left(\frac{\mathbf{C}_P}{|\mathbf{C}_P|} \cdot \frac{\mathbf{C}_E}{|\mathbf{C}_E|}\right). \tag{5}$$

Here, '$\cdot$' denotes the inner product operator. For a more detailed coverage of PCC we refer readers to [11].

III. PCC Application I: Google Orkut

In this section we apply PCC to a large scale data set obtained from the friend network of the Orkut social networking service [12]. The first data set is a friendship graph of $70,000$ users of Google's Orkut social network service. This data set was originally collected by Mislove [14] and constitutes an unweighted, undirected graph. The data set is a social graph obtained from subscriber friends lists of Google's Orkut social network service [14]. The data set consists of $70,000$ user nodes with $2,971,776$ undirected links between them and was processed and analyzed on Matlab 7.4 (R2007a) on a Dell PowerEdge server with an Intel QuadCore Xeon $2.13GHz$ processor and $4GB$ RAM.

Our objective for applying PCC on this data set is the discovery of more social hubs than are identified by EVC. Figure 1a plots the EVC scores of all $70,000$ nodes in the data set. It shows that node 692 in the network has the highest EVC, followed by a cluster of nodes with node IDs centered around $43,000$. The remaining majority of nodes is assigned centrality scores close to 0. The histogram of node EVCs in figure 1b confirms this. The number of features P to use for the computation of PCC is the number of eigenvalues after which the rate of growth of their cumulative sum begins to decline significantly. An alternative approach which yields a clearer cutoff point for the selection of the number of features to use in PCC was the plot of the rate of change in the phase angle of PCC vectors with the EVC vector. Figure 1c plots the phase angles ϕ of PCC while varying the number of features from 1 through 100. We select $P = 14$ as a cut-off point for the computation of PCC (marked in red). Figure 1e plots PCC vector of all nodes in the Orkut graph. When we compare it to the plot of the EVC vector what stands out immediately is how a lot of nodes with near-zero EVCs are assigned higher and highly varying PCC scores. Figure 1f is the histogram of PCC scores which, at a standard deviation of 6839.7, is more spread out than that of the EVC with a standard deviation of 4382.7.

Figure 1d plots the number of local maxima that are found in the graph for values of P. At first glance it might appear that EVC found 84, while PCC improves this number to 91 when using just 14 out of $70,000$ possible features. In the plotted range a maximum of 95 maxima are identified when 61 features are used. We examine how many of the social hubs identified are in fact trivial maxima with low centrality. This is accomplished by viewing the centrality scores of nodes identified as social hubs. Let S_n denote the set of social hubs identified by using PCC with n features/eigenvectors. Figure 1g is the histogram of EVC histogram of the top 20 EVC scoring nodes in S_1 (set of social hubs identified using only EVC/PCC with 1 feature). Only one node (node number 692, the node with the highest EVC of all $70,000$ nodes in figure 1b) truly stands out with a high EVC, whereas the other 19 social hubs' EVC scores lie in the lowest bin of the histogram. Thus, after the definition of local maxima excludes nodes surrounding the most central node, EVC fails to identify any other influential neighborhood. One might wonder if, among the more than 2000 nodes with pronounced EVC scores in figure 1b there is not a single node besides node 692 that might be a local maxima. We verified from the data set that node 692 has 2185 neighbors, most of which have node IDs in the range between 42134 and 44314 (clearly visible in figure 1b). In contrast, figure 1h is a histogram of the PCC scores of the 20 nodes with highest PCC scores in S_{14} (the set of social hubs identified using $\mathbf{C}_{14}$, PCC with 14 features). Here, a total of 8 social hubs have non-trivial PCC scores, the remaining 12 social hubs have PCC scores too low to be considered significant. The IDs of nodes identified as social hubs of influential neighborhoods in descending order of PCCs are 692, 317, 4749, 487, 39, 14857, 35348 and 12219. This is a substantial improvement over the single neighborhood identified using EVC. Thus, using PCC in conjunction with a node selection criteria provided by the definition of local maxima identifies many more influential neighborhoods in a social network than is possible by using EVC.

We also raise the question of how different the set of nodes identified as social hubs is from the nodes we would have identified as central were we to rely solely on nodes' centrality scores. This raises the question of how different S_P, the set of nodes identified as social hubs based on $\mathbf{C}_P$, is from $V(\mathbf{C}_P[|S_P|])$, the vertex set (returned by the function $V()$) of the first $|S_P|$ nodes ranked in descending order of $\mathbf{C}_P$. Figure 1i plots the size of the intersection set $|S_P \cap V(\mathbf{C}_P[|S_P|])|$. The data point at $P = 1$ is 1 and is the number of nodes common in the set of social hubs identified by EVC and those identified by node EVC scores alone. For the range $1 \leq P \leq 100$ at most 3 nodes identified as social hubs are also present in the set of the first $|S_P|$ most central nodes, i.e. placing monitors at nodes based solely on their centrality scores produces a lot of redundant coverage.

We compute the sizes of intersection sets of all pairs of S_{P1} and S_{P2} for $1 \leq P1, P2 \leq 100$. This is plotted in figure 1j. It shows that as the number of features $P1$ used to compute $\mathbf{C}_{P1}$ is increased from $\mathbf{C}_{P2}$, the set of social hubs S_{P2} identified by it is (almost) always a superset of the set S_{P1} if $P1 < P2$. Thus the inclusion of more feature vectors adds members to the set of social hubs without removing previous ones.

IV. PCC Application II: Facebook Fighter's Club App

The second data set is derived from a list of matches played between users of Facebook's 'Fighters' Club' application. This data set contains $143,020$ users and $667,560$ matches was originally collected by Nazir in [15]. It differs from the Orkut friendship graph in that it is a weighted, undirected graph. The weights of links between two user nodes represent the number of interactions/matches played against each other. Each vertex in the weighted gaming graph represents a user of the application with $526,224$ edges between them. Thus, weight of an edge between two users is the number of matches recorded between them in the data set. Edge weights in this data set range from 1 to 29.

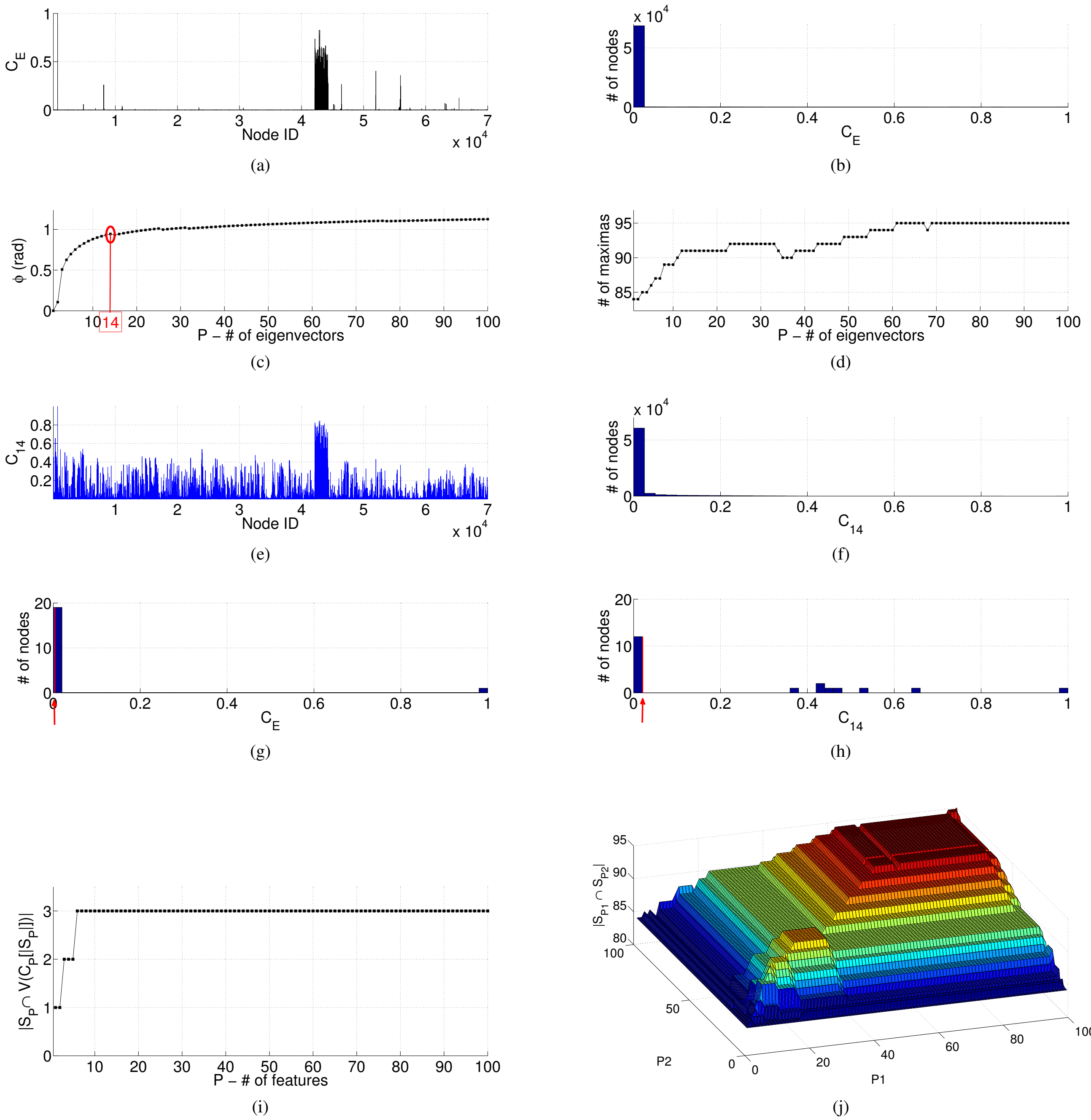

Fig. 1. Orkut data set: a) EVC scores of nodes, b) A histogram of EVC scores, c) Phase angles ϕ of PCC vectors $\mathbf{C}_1$ through $\mathbf{C}_{100}$ with EVC vector $\mathbf{C}_E$, d) Number of local maxima discovered using PCC's of varying number of features, e) PCC scores of nodes using 14 features, f) A histogram of PCC scores of nodes using 14 features, g) A histogram of EVCs of the 20 social hubs with the highest EVCs, and h) A histogram of PCCs of the 20 social hubs with the highest PCCs based on 14 features, i) The size of the intersection set of S_P and $V(\mathbf{C}_P[|S_P|])$ for $1 \leq P \leq 100$, and j) The size of the intersection set of S_{P1} and S_{P2} when $1 \leq P1, P2 \leq 100$.

Figure 2a plots the EVC scores of all $143,020$ nodes in the data set. Unlike in the preceding Orkut data set, there are only very few nodes that are assigned EVCs significantly greater than 0. The group of nodes in the node ID space above $130,000$ are the only ones with high EVCs, while almost all other nodes have near-zero EVCs. Like in the example illustrated earlier, the remaining majority of nodes is assigned centrality scores close to 0. The histogram of node EVCs in figure 2b confirms this. Figure 2c plots the phase angles ϕ of the PCC vector while varying the number of features from 1 through 100. Figure 2d plots the number of local maximas that are found in the gaming graph for values of $1 \leq P \leq 100$. EVC finds approximately 1.23×10^5 while PCC increases this number slightly to 1.232×10^5 when using just 20 out of

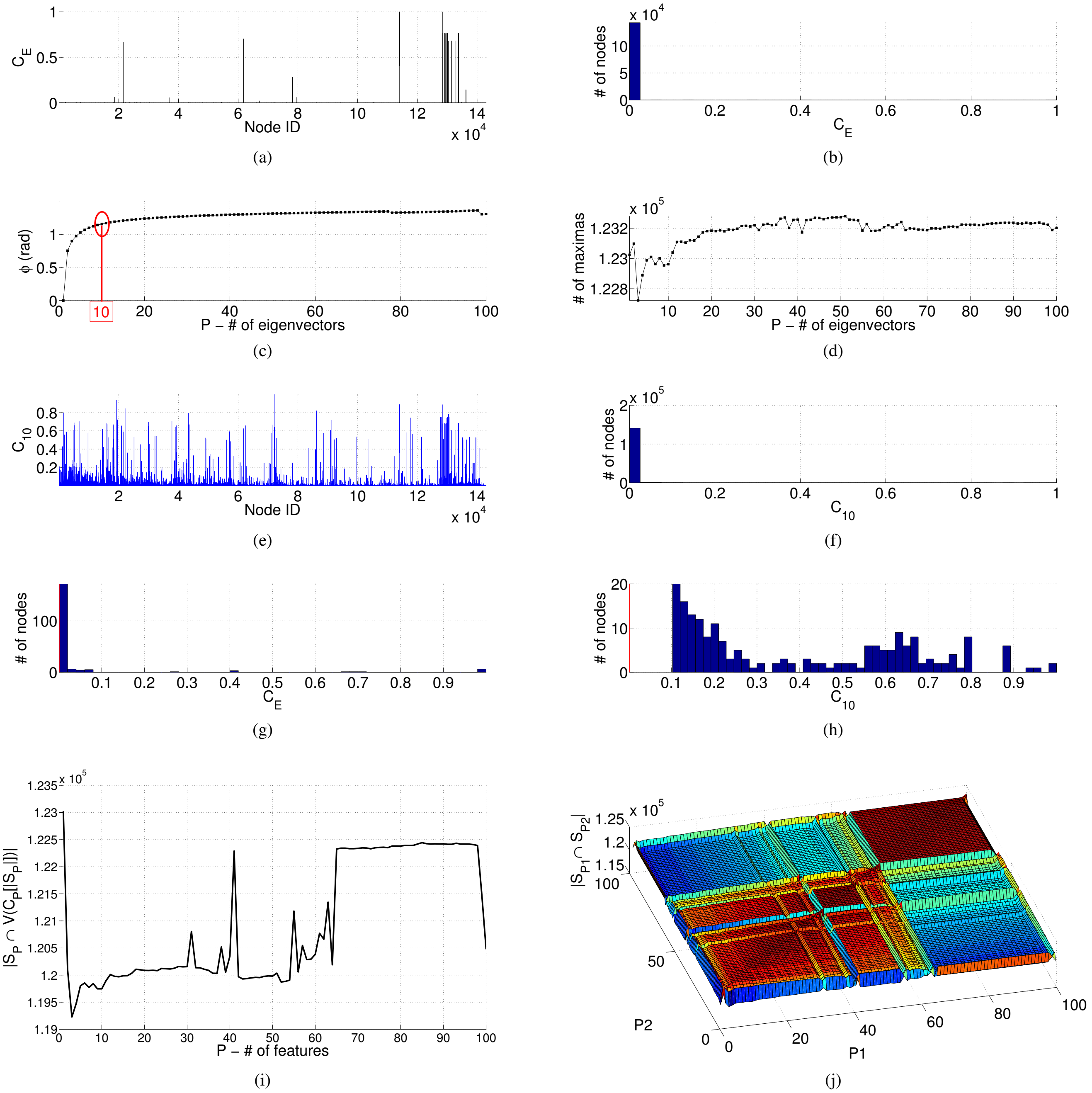

Fig. 2. Facebook Fighters' Club application data set: a) EVC scores of nodes, b) A histogram of EVC scores, c) Phase angles ϕ of PCC vectors $\mathbf{C}_1$ through $\mathbf{C}_{100}$ with EVC vector $\mathbf{C}_E$, d) Number of local maxima discovered using PCC's of varying number of features, e) PCC scores of nodes using 10 features, f) A histogram of PCC scores of nodes using 10 features, g) A histogram of EVCs of the 200 social hubs with the highest EVCs, and h) A histogram of PCCs of the 200 social hubs with the highest PCCs based on 10 features, i) The size of the intersection set of S_P and $V(\mathbf{C}_P[|S_P|])$ for $1 \leq P \leq 100$, and j) The size of the intersection set of S_{P1} and S_{P2} when $1 \leq P1, P2 \leq 100$.

$143,020$ possible features. The phase angle ϕ attains a stable value around $P = 10$ features (marked by red line) and so we will use $P = 10$ for the computation of PCC, i.e. $\mathbf{C}_{10}$. Figure 2e plots PCCs of all nodes in the graph with 10 features and figure 2f is their histogram.

This raises the question of how different S_P, the set of nodes classified as social hubs based on $\mathbf{C}_P$, is from $V(\mathbf{C}_P[|S_P|])$, the vertex set (returned by the function $V()$) of the first $|S_P|$ nodes ranked in descending order of $\mathbf{C}_P$. Figure 2i plots the size of the intersection set $|S_P \cap V(\mathbf{C}_P[|S_P|])|$. The data point at $P = 1$ is approximately 1.23×10^5 and is the number of nodes common in the set of social hubs identified by EVC and those identified by node EVC scores alone. As we proceed on the horizontal axis the number of features used to compute PCC is increased. Each data point is the number of nodes in the intersection of the set of social hubs by P-feature PCC

($\mathbf{C}_P$) with the set of most central nodes nodes by PCC $\mathbf{C}_P$ of equal size. As P increases in the range $1 \leq P \leq 100$ the size of the intersection set rapidly drops at first and then climbs back close to the starting value at 65 features.

We compute the sizes of intersection sets of all pairs of S_{P1} and S_{P2} for $1 \leq P1, P2 \leq 100$. This is plotted in figure 2j. It shows that as the number of features $P1$ used to compute $\mathbf{C}_{P1}$ is increased from $\mathbf{C}_{P2}$, the set of social hubs S_{P2} identified by it is (almost) always a superset of the set S_{P1} if $P1 < P2$. Thus the inclusion of more feature vectors retains a large fraction of social hubs identified using fewer features.

V. Conclusions

We reviewed previously defined measures of centrality and pointed out their shortcomings in general and EVC in particular. We reviewed PCC, a new measure of node centrality. PCC is based on PCA which takes the view of treating a graphs adjacency matrix as a covariance matrix. PCC interprets a node's centrality as its ℓ^2 norm from the origin in the eigenspace formed by the P most significant feature vectors (eigenvectors) of the adjacency matrix. Unlike EVC, PCC allows the addition of more features for the computation of node centralities. We explore two criteria for the selection of the number of features to use for PCC; a) The relative contribution of each feature's power (eigenvalue) to the total power of adjacency matrix and b) Incremental changes in the phase angle of the PCC with P features and the EVC as P is increased.

We applied PCC analysis to Google's Orkut social networking service. Our objective was the identification of social hubs in social networks that are left undiscovered by EVC. In the case of the Orkut graph we saw that using 14 most significant eigenvectors out of a possible $70,000$ raises the number of influential neighborhoods discovered from just 1 (that around the most central node) to 8 (including the one identified by EVC). The increase in the number of social hubs found using PCC is even greater. The top 200 social hubs found using PCC all have normalized PCC greater than 0.1. The social hubs found using EVC however contain only 13 social hubs with normalized EVC greater than 0.1.

The Orkut friendship graph we used was unweighted and undirected, while the Facebook application graph was an undirected and weighted graph. However, in order to ensure that eigenvalues and eigenvectors remain real the graphs must be undirected. We compared the sets of nodes identified as social hubs with the set of highest scoring nodes by centrality alone and saw that many of the nodes with high centralities belong to the same neighborhood. Thus, the notion of a local maxima serves its purpose of removing neighbors of highly central nodes. A comparison of social hubs discovered by PCC with different numbers of features showed that the addition of more features in PCC adds new social hubs to the list of identified hubs without replacing previously identified ones.

In the future we intend to extend the definition of PCC so it can be applied to both directed and undirected graphs. Furthermore, we propose to formulate a distributed method for computing PCC along the lines of Canright's method [5] for computing EVC in peer-to-peer systems.

References

[1] P. Bonacich, "Power and centrality: A family of measures," *American Journal of Sociology*, vol. 92, pp. 1170–1182, 1987.

[2] P. Bonacich and P. Lloyd, "Eigenvector-like measures of centrality for asymmetric relations," *Social Networks*, vol. 23, no. 3, pp. 191–201, 2001.

[3] S. Borgatti, "Centrality and network flow," *Social Networks*, vol. 27, no. 1, pp. 55–71, 2005.

[4] S. Borgatti and M. Everett, "A graph-theoretic perspective on centrality," *Social Networks*, vol. 28, no. 4, pp. 466–484, 2006.

[5] G. Canright, K. Engø-Monsen, and M. Jelasity, "Efficient and robust fully distributed power method with an application to link analysis," *Department of Computer Science, University of Bologna, Tech. Rep. UBLCS-2005-17*, pp. 2005–17, 2005.

[6] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. Wiley New York, 2001.

[7] M. Everett and S. Borgatti, *Extending centrality*, P. J. Carrington, J. Scott, and S. Wasserman, Eds. Cambridge University Press, 2005.

[8] L. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.

[9] N. Friedkin, "Theoretical foundations for centrality measures," *American Journal of Sociology*, vol. 96, no. 6, p. 1478, 1991.

[10] H. He, "Eigenvectors and reconstruction," *the electronic journal of combinatorics*, vol. 14, no. 14, p. 1, 2007.

[11] M. Ilyas and H. Radha, "A klt-inspired node centrality for identifying influential neighborhoods in graphs," in *Proceedings of the 44th International Conference on Information sciences and systems*, March 17-19 2010, pp. 1 – 7.

[12] G. Inc., "orkut - welcome to the orkut club!" http://www.orkut.com/Signup, Tech. Rep., Oct 14, 2009.

[13] A. Langville, C. Meyer, and P. FernÁndez, "Googles pagerank and beyond: the science of search engine rankings," *The Mathematical Intelligencer*, vol. 30, no. 1, pp. 68–69, 2008.

[14] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proceedings of the 5th ACM/Usenix Internet Measurement Conference (IMC'07)*, San Diego, CA, October 2007.

[15] A. Nazir, S. Raza, and C. Chuah, "Unveiling facebook: a measurement study of social network based applications," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*. ACM New York, NY, USA, 2008, pp. 43–56.

[16] Orkut, "Site profile for orkut.com (rank 4,168) — complete," http://siteanalytics.compete.com/orkut.com/?metric=uv, Tech. Rep., September, 2009.

[17] B. Ruhnau, "Eigenvector-centrality a node-centrality?" *Social Networks*, vol. 22, no. 4, pp. 357–365, 2000.

[18] F. Tisseur and J. Dongarra, "A parallel divide and conquer algorithm for the symmetric eigenvalue problem on distributed memory architectures," *SIAM Journal on Scientific Computing*, vol. 20, p. 2223, 1999.

[19] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*. Cambridge University Press, 1994.

[20] D. Watts, "Networks, dynamics, and the small-world phenomenon," *American Journal of Sociology*, vol. 105, no. 2, pp. 493–527, 1999.