IJCSS International Journal of Complex Systems in Science

Int. J. Complex Systems in Science
vol. **3(1)** (2013), pp. 113–119

Competitivity graphs analysis and structural comparison of rankings

Regino Criado[1,2], Esther García[1,2], Francisco Pedroche[3,†], Miguel Romance[1,2] and Victoria E. Sánchez-García[4]

[1] *Departamento de Matemática Aplicada, Universidad Rey Juan Carlos, Spain*

[2] *Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Spain*

[3] *Institut de Matemàtica Multidisciplinària, Universitat Politècnica de València*

[4] *Departamento de Economía de la Empresa, Universidad Carlos III, Spain*

Abstract. A complex networks based method is introduced for comparing different complete rankings of a finite family of elements. The concepts of *competitivity graph* and *evolutive competitivity graph* are introduced as the main tools for analyzing an (ordered) family of rankings of a fixed set of elements. It is shown how the structural properties of these competitivity graphs give deep information about the competitiveness of the elements according to the rankings considered. The relationships between competitivity graphs and some other well-known families of graphs, such as permutation graphs, comparability graphs and chordal graphs are also presented. Finally some applications are presented, including the analysis of sports rankings and, more precisely, the study of European soccer leagues.

Keywords: Ranking analysis, Competitiveness, Complex network analysis, Comparability graphs.
MSC 2000: 05C75, 05C82

[†] **Corresponding author:** pedroche@mat.upv.es
Received: December 5th, 2013
Published: December 31th, 2013

1. Introduction

Complex networks have been the subject of intense study in recent years: Internet, the World Wide Web, and many other types of technological, biological and social networks have given us a new insight of the possibilities of this new branch of science. Complex networks are objects composed by a set of nodes or vertices that are pairwise joined by links or edges. As we have said, this kind of representation has been recently and successfully used in various technological, social and biological scenarios, but the study of networks has a long history in mathematics, inside of a branch of discrete mathematics known as graph theory. The main difference between the network theory's approach and the (combinatorial) graph theory's approach is that the analysis of complex network always takes care of the computational complexity of the studied problems, due to the big (generally huge) number of nodes of a network.

Rankings are everywhere. We know different types of rankings that classify universities by their prestige, countries by their gross domestic product (GDP), companies by the price and the evolution of their stocks, sportsmen by their marks,... So it is necessary to develop tools and improve them in order to analyze in depth and compare the real situation and the evolution of many kind of objects that compete amongst them, being reflected the result of that competition through a family of rankings.

The comparison of families of rankings has been a topic of interest of several authors. We highlight the seminal work of Kendal [8], where the Kendall's concordance coefficient is defined. The Kendall's correlation coefficient τ for two rankings was introduced in his previous work [7]. Rankings can be also compared by measuring their distance, for example, by the Spearman's footrule rule or some other metrics, see [10], [5]. The number of papers dealing with ranking systems is huge, including those describing a graph to define a ranking. Nevertheless, the novelty of our work is the use of graphs as a tool to compare families of rankings.

2. Competitivity Graphs

Given a set of elements $\mathcal{N} = \{1, \ldots, n\}$ that we will call *nodes* we define a *ranking* c of $\mathcal{N}$ as any bijection $c : \mathcal{N} \to \mathcal{N}$. We will identify rankings with vectors of $\mathcal{N}^n$ in the following way: $c \equiv (i_1, \ldots, i_n)$ if $c(1) = i_1, c(2) = i_2, \ldots,$ $c(n) = i_n$. We will write $i \prec_c j$ when node i appears first than node j in the vector of the ranking c, i.e., when $c(i) < c(j)$.

Given a finite set $\mathcal{R} = \{c_1, c_2, \ldots, c_r\}$ of rankings we say that the pair of nodes $(i, j) \in \mathcal{N}$ *compete* if there exists $c_s, c_t \in \{1, 2, \ldots, r\}$ such that $i \prec_{c_s} j$ but $j \prec_{c_t} i$, i.e., i and j exchange their relative positions between the rankings

c_m and c_n. We define the *competitivity graph* of the family of rankings $\mathcal{R}$, denoted by $G_c(\mathcal{R}) = (\mathcal{N}, E_{\mathcal{R}})$, where $E_{\mathcal{R}}$ denotes the set of edges, as the undirected graph with nodes $\mathcal{N}$ and edges given by the rule: there is a link between i and j if (i, j) compete.

When the family of rankings $\mathcal{R} = \{c_1, c_2, \dots, c_r\}$ is ordered, we say that a pair of nodes $(i, j) \in \mathcal{N}$ *compete at raking* c_s if they exchange their relative positions between the rankings c_s and c_{s+1}. We say that two nodes i, j compete k-times if k is the maximal number of rankings where i and j compete. The *evolutive competitivity graph* of $\mathcal{R}$, denoted by $G_c^e(\mathcal{R}) = (\mathcal{N}, E_{\mathcal{R}}^e)$, will be the weighted undirected graph with nodes $\mathcal{N}$ and edges given by the rule: there is an edge between i and j labeled with weight k if (i, j) compete k times. Note that the underlying (unweighted) network behind the (weighted) graph $G_c^e(\mathcal{R})$ is $G_c(\mathcal{R})$.

Note that the order of the rankings is fundamental in the calculation of the weights of the evolutive competitivity graph, although it does not have influence in the underlying (unweighted) competitivity graph.

Competitivity graphs have already been studied in the particular case of two rankings ($r = 2$). They are the so-called *permutation graphs*, see [4]. Permutation graphs are a subclass of another classical class of graphs: comparability graphs, see [6], [9]. In the following results of [2] we relate competitivity graphs with comparability graphs, and also with the class of chordal graphs, see [3].

Competitivity versus comparability:

(i) There are comparability graphs that are not competitivity graphs.

(ii) There are competitivity graphs that are not comparability graphs.

(iii) There are graphs that are neither comparability nor competitivity graphs.

(iv) Permutation graphs are both competitivity and comparability graphs.

Competitivity versus chordal:

(i) There are chordal graphs that are not competitivity graphs.

(ii) There are competitivity graphs that are not chordal graphs.

(iii) There are graphs that are neither competitivity nor chordal.

(iv) There are graphs that are both competitivity and chordal.

A deeper study of the structural properties of competitivity graphs has also been done in [2]. We highlight the computation of the set of eventual competitors (connected components of the competitivity graph) directly from the rankings without the previous computation of the competitivity graph. For example, if the graph has more than one set of eventual competitors, the elements of $\mathcal{N}$ can be separated into subsets of elements that never compete among them.

3. Some applications

There are several ways to define the *competitiveness* of two or more ordered families of rankings $\mathcal{R} = \{c_1, c_2, \ldots, c_r\}$ and $\mathcal{S} = \{c'_1, c'_2, \ldots, c'_s\}$ possibly coming from different sets of nodes or competitors $\mathcal{N} = \{1, \ldots, n\}$ and $\mathcal{N}' = \{1, \ldots, n'\}$. Let $G_c^e(\mathcal{R}) = (\mathcal{N}, E_{\mathcal{R}}^e)$ and $G_c^e(\mathcal{S}) = (\mathcal{N}', E_{\mathcal{S}}^e)$ be two different evolutive competitivity graphs. The underlying unweighted competitivity graphs will be denoted by $G_c(\mathcal{R}) = (\mathcal{N}, E_{\mathcal{R}})$ and $G_c(\mathcal{S}) = (\mathcal{N}', E_{\mathcal{S}})$. As measures of competitiveness we will consider different parameters:

Normalized mean degree. We define the *normalized mean degree* of a family of rankings $\mathcal{R}$ as the sum of all the node degrees in the competitivity graph $G_c(\mathcal{R})$ divided by the sum over all nodes of their highest possible degrees

$$\mathrm{ND}(\mathcal{R}) = \frac{1}{n(n-1)} \sum_{i \in \mathcal{N}} \deg(i). \tag{1}$$

We say that $\mathcal{R}$ *is more competitive than* $\mathcal{S}$ *with respect to the normalized mean degree* if $\mathrm{ND}(\mathcal{R}) > \mathrm{ND}(\mathcal{S})$.

Normalized mean strength. The strength of a node in a weighted graph is the sum of the weights of its incident edges. We define the *normalized mean strength* of a family of rankings $\mathcal{R}$ as the sum of all edge weights in the evolutive competitivity graph $G_c^e(\mathcal{R})$ divided by the sum over all possible edges of their highest possible weights:

$$\mathrm{NS}(\mathcal{R}) = \frac{w(E_{\mathcal{R}}^e)}{\binom{n}{2}(r-1)}, \tag{2}$$

where $w(E_{\mathcal{R}}^e)$ denotes the sum of all weights of the edges of the evolutive competitivity graph.

We say that $\mathcal{R}$ *is more competitive than* $\mathcal{S}$ *with respect to the normalized mean strength* if $\mathrm{NS}(\mathcal{R}) > \mathrm{NS}(\mathcal{S})$.

Clustering coefficient. In graph theory, a clique is a set of nodes mutually connected between them. For example, a triangle is a clique formed by three nodes. The clustering coefficient measures how many nodes in a graph tend to cluster together. The clustering coefficient C_i of a node i is defined as

$$C_i = \frac{e_i}{\binom{k_i}{2}}, \tag{3}$$

where k_i is the number of neighbors of node i, e_i is the number of connected pairs between the neighbors of i, and $\binom{k_i}{2}$ represents all possible pairs between the neighbors of i. Given a family of rankings $\mathcal{R}$, the *clustering coefficient* of $\mathcal{R}$ is the average of the clustering coefficients of the nodes of the competitivity graph $G_c(\mathcal{R})$, i.e.,

$$C(\mathcal{R}) = \frac{1}{n} \sum_{i \in \mathcal{N}} C_i. \tag{4}$$

We say that $\mathcal{R}$ *is more competitive than* $\mathcal{S}$ *with respect to the clustering coefficient* $\mathcal{C}$ if $\mathcal{C}(\mathcal{R}) > \mathcal{C}(\mathcal{S})$.

Similarly, we can consider other graph parameters such as the normalized size of the maximal clique (i.e., the number of nodes of the maximal clique contained in the graph divided by the number of nodes of the graph) the normalized size of the largest connected component, etc. For each of these parameters, a family of rankings is more competitive than another family if this parameter in the (evolutive) competitivity graph is bigger

There are other graph parameters that work the other way round: the smaller they are, the more competitive a family of rankings is. Examples of such parameters are the number of connected components and the Kendall's coefficient τ:

Generalized Kendall's τ correlation coefficient. We can define a *generalized Kendall's correlation coefficient* $\tau(\mathcal{R})$ of a family $\mathcal{R}$ of $r \geq 2$ rankings. Following the original definition (number of pairs that do not compete $\tilde{K}(\mathcal{R})$ minus number of pairs that compete $K(\mathcal{R})$, divided by the number of all possible pairs $\binom{n}{2}$, see [7]), we set

$$\tau(\mathcal{R}) = \frac{\tilde{K}(\mathcal{R}) - K(\mathcal{R})}{\binom{n}{2}} = 1 - \frac{2|E_{\mathcal{R}}|}{\binom{n}{2}} = 1 - \frac{4|E_{\mathcal{R}}|}{n(n-1)}$$

where $|E_{\mathcal{R}}|$ denotes the number of edges of the competitivity graph $G_c(\mathcal{R})$.

We can also construct an *evolutive Kendall's correlation coefficient* $\tau(\mathcal{R})_e$ if we take into account the number of times each pair of nodes compete. In this sense, we define

$$\tau(\mathcal{R})_e = 1 - \frac{2\ w(E^e_{\mathcal{R}})}{\binom{n}{2}(r-1)}, \tag{5}$$

where $w(E^e_{\mathcal{R}})$ denotes the sum of all weights of the edges of the evolutive competitivity graph. The denominator $\binom{n}{2}(r-1)$ represents the sum over all possible edges of their highest possible weights.

We say that $\mathcal{R}$ *is more competitive than* $\mathcal{S}$ *with respect to the Kendall's coefficient* if $\tau_e(\mathcal{R}) < \tau_e(\mathcal{S})$. Notice that the smaller the Kendall's coefficient $\tau_e(\mathcal{R})$ is, the more competitive $\mathcal{R}$ is.

By using these parameters we have compared the competitiveness of the mayor European soccer leagues in 2011-12 and 2012-13, see [1]. For example, Figure 1 shows the evolution of the normalized mean strength along the seasons 2011-12 (on the left) and 2012-13 (on the right).

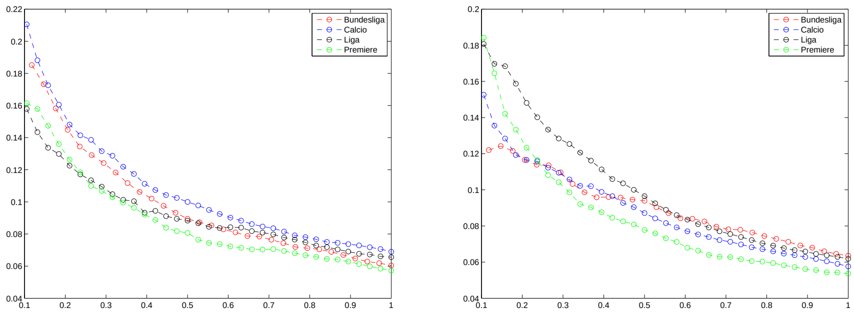

Figure 1: The evolution of the normalized mean strength along 2011-12 (on the left) and 2012-13 (on the right). In both figures the German Bundesliga is in red, the Italian Lega A is in blue, the Spanish Liga BBVA is in black and the British Premier League is in green.

4. Conclusions

In this paper we have presented a new tool for analyzing and comparing complete rankings of a finite family of elements. Two new concepts within the realm of graph's theory are defined: the competitivity graph of a family of rankings, and the evolutive competitivity graph of an ordered family of rankings. The structural properties of these graphs give deep information about the competitiveness of the elements according to the rankings considered.

We highlight that this methodology can be used to study not only sport rankings but other families of (ordered) rankings: evolution of stock markets, classifications of countries with respect to different parameters, ordered lists of universities, etc.

Acknowledgements

This work was partially supported by Spanish MICINN Funds and FEDER Funds MTM2009-13848, MTM2010-16153 and MTM2010-18674, and Junta de Andalucía Funds FQM-264.

References

[1] R. Criado, E. García, F. Pedroche and M. Romance, *A new method for comparing rankings through complex networks: Model and analysis of competitiveness of major European soccer leagues*, Chaos **23**, 033131 (2013).

[2] R. Criado, E. García, F. Pedroche and M. Romance, *Comparing rankings by means of competitivity graphs: structural properties and computation*, arXiv:1310.6921 (2013), 1-12.

[3] G. A. Dirac, *On rigid circuit graphs*, Abh. Math. Sem. Univ. Hamburg **25** (1961), 71–76.

[4] B. Dushnik and E.W Miller, *Partially ordered sets*, Amer. J. Math. **63** (1941), 600-610.

[5] R. Fagin, R. Kumar, M. Mahdian, D. Sivakumar and E.Vee, *Comparing Partial Rankings*, SIAM J. Discrete Math. **20**(3) (2006), 628-648.

[6] A. Ghouila-Houri *Caractérisation des graphes non orientés dont on peut orienter les arrêtes de manière à obtenir le graphe d'une relation d'ordre*, C. R. Acad. Sci. Paris **254** (1962), 1370-1371.

[7] M. G. Kendall, *A New Measure of Rank Correlation*, Biometrika **30**(1-2) (1938), 81-89.

[8] M. G. Kendall and B. B. Smith, *The Problem of m Rankings*, Ann. Math. Statist. **10**(3) (1939), 275-287.

[9] A. Pnueli, A. Lempel and S. Even, *Transitive Orientation of Graphs and Identification of Permutation Graphs*, Canad. J. Math. **23** (1) (1971), 160-175.

[10] P. K. Sen, I. A. Salama and D. Quade, *Spearman's footrule: Asymptotics in applications* Chilean J. Stat. **2**(1) (2011), 3-20.