# Centrality in Dynamic Competition Networks

Anthony Bonato[1(✉)], Nicole Eikmeier[2], David F. Gleich[3], and Rehan Malik[1]

[1] Ryerson University, Toronto, ON, Canada
abonato@ryerson.ca
[2] Grinnell College, Grinnell, IA, USA
[3] Purdue University, West Lafayette, IN, USA

Abstract. Competition networks are formed via adversarial interactions between actors. The Dynamic Competition Hypothesis predicts that influential actors in competition networks should have a large number of common out-neighbors with many other nodes. We empirically study this idea as a centrality score and find the measure predictive of importance in several real-world networks including food webs, conflict networks, and voting data from Survivor.

1 Introduction

While social networks are often studied from the perspective of positive interactions such as friendship or followers, the impact of negative social interaction on their structure and evolution cannot be ignored. Structural balance theory posits positive and negative ties between actors in social networks, and assumes such signed networks will stabilize so that triples of actors are either all mutually friends or possess common adversaries; see [12], and [9] for a modern treatment. The prediction of the signs of edges in a social network was previously studied [15,18,21]. Further, negative interactions as a model for edges was studied in the context of negatively correlated stocks in market graphs [4], and in the spatial location of cities as a model to predict the rise of conflicts and violence [11]. Even in the highly cited Zachary Karate club network [22], the negative interaction between the administrator and instructor was the impetus for the split of the club participants into two communities. We propose that competition or negative interactions are critically important to the study of social networks and more broadly, real-world complex networks, and are often hidden drivers of link formation.

In [6], we investigated properties inherent in social networks of competitors that evolve dynamically over time. Such networks are viewed as directed, where a directed edge from nodes u to v corresponds to some kind of negative social interaction. For example, a directed edge may represent a vote by one player for another in a social game such as the television program Survivor. Directed edges are added over discrete time-steps in what we call dynamic competition networks. Our main contribution in [6] was the presentation of a hypothesis, referred to as the Dynamic Competition Hypothesis, or (DCH), that served as a predictive tool to uncover alliances and leaders within dynamic competition

networks. We provided evidence for the hypothesis using U.S. voting record data from 35 seasons of Survivor.

In the present paper, we focus on a particular implication of the DCH. Namely, the DCH predicts that leaders and central actors in these networks should have large a large number of common out-neighbors with other nodes in the network. Consequently, this score should constitute a more accurate and interesting centrality score in competition networks where edges have a negative connotation. We study this score in terms of its ranking of leaders in various kinds of networks ranging from additional international seasons of Survivor, to conflict networks, and to food webs.

We organize the discussion in this paper as follows. In Sect. 2, we formally define dynamic competition networks, and review the DCH as stated in [6], with a focus on the common out-neighbor scores, called CON scores. In Sect. 3, we investigate using CON scores as centrality measures in three distinct sources: (i) voting data from all international (that is, non-U.S.) seasons of Survivor, (ii) from conflict networks arising from the countries of Afghanistan, India, and Pakistan, and (iii) in 14 food webs. We find that the CON scores predict influential actors in the networks with high precision. The final section interprets our results for real-world complex networks, and suggests further directions.

We consider directed graphs with multiple directed edges throughout the paper. Additional background on graph theory and complex networks may be found in the book [5] or [7].

2   The Dynamic Competition Hypothesis

The Dynamic Competition Hypothesis (DCH) provides a quantitative framework for the structure of dynamic competition networks. We recall the statement of the DCH as first stated in [6]. Before we state the DCH, we present some terminology.

A *competition network* G is one where nodes represent actors, and there is directed edge between nodes u and v in G if actor u is in competition with actor v. A *dynamic competition network* is a competition network where directed edges are added over discrete time-steps. For example, nodes may consist of individuals and edges correspond to conflicts between them; as another example, we may consider species in an ecological community, and directed edges correspond to predation. Observe that dynamic competition networks may have multiple edges if there were multiple conflicts; further, not all edges need be present.

The central piece of the DCH we study here are the *common out-neighbor* scores. Without loss of generality, we assume that the node correspond to integers such that we can use the nodes to address an adjacency matrix as well. Consequently, let $\boldsymbol{A}$ be the adjacency matrix of given competition network. Entries in the matrix are 0 or positive integers for the number of competition interactions. For nodes $u, v,$ and w, we say that w is a *common out-neighbor* of u and v if (u, w) and (v, w) are directed edges. Alternatively, $A_{uv}A_{vw} \geq 1$. For a pair of distinct nodes u, v, we define $\text{CON}(u, v)$ to be the number of common out-neighbors of u and v. Note that this common out-neighbor score counts multiplicities based

on the minimum number of interactions: $\text{CON}(u, v) = \sum_k \min(A_{uk}, A_{vk})$, which corresponds to a multiset intersection. For a fixed node u, define

$$\text{CON}(u) = \sum_{v \in V(G)} \text{CON}(u, v).$$

We call $\text{CON}(u)$ the *CON score* of u. For a set of nodes S with at least two nodes, we define

$$\text{CON}(S) = \sum_{u, v \in S} \text{CON}(u, v).$$

Observe that $\text{CON}(S)$ is a non-negative integer.

In the DCH, *leaders* are defined as members of a competition network with high standing in the network, and edges emanating from leaders may influence edge creation in other actors. In the context of conflict networks within a country, leaders may be actors who exert the strongest political influence within the country; note that these may not be the largest or most powerful actors. As another example, leaders in a food web would naturally have higher trophic levels (that is, higher position in a food chain). The DCH characterizes leaders as those nodes with high CON scores, low in-degree, high out-degree and high closeness. Recall that for a strongly connected digraph G and a node v, we define the *closeness* of u by

$$C(u) = \left(\sum_{v \in V(G) \setminus \{u\}} d(u, v) \right)^{-1}$$

where $d(u, v)$ corresponds to the distance measured by one-way, directed paths from u to v.

In this paper, we focus on the implication that leaders in competition networks should have high CON scores, which suggests this is a natural centrality measure for these networks. The DCH also involves the notion of alliances, that does not factor into our present study. *Alliances* are defined as groups of agents who pool capital towards mutual goals. In the context of social game shows such as Survivor, alliances are groups of players who work together to vote off players outside the alliance. Members of an alliance are typically less likely to vote for each other, and this is the case in strong alliances. This is characterized in terms of *near independent sets*; see [6] for the formalism.

In summary, the *Dynamic Competition Hypothesis* (or *DCH*) asserts that dynamic competition networks satisfy the following four properties.

1. Alliances are near independent sets.
2. Strong alliances have low edge density.
3. Members of an alliance with high CON scores are more likely leaders.
4. Leaders exhibit high closeness, high CON scores, low in-degree, and high out-degree.

Our focus in this work will be on the validation of the DCH with regards to detecting leaders; in particular, we will focus on items (3) and (4) of the DCH. Note that while we expect leaders to be in alliances (that is, have prominent local influence), leaders are determined via global metrics of the network.

3   Methods and Data

3.1   Survivor

In [6], we studied the voting history of U.S. seasons of Survivor, which is a social game show where players compete by voting each other out. In Survivor, strangers called survivors are placed in a location and forced to provide shelter and food for themselves, with limited support from the outside world. Survivors are split into two or more tribes which cohabit and work together. Tribes compete for immunity and the losing tribe goes to tribal council where one of their members is voted off. At some point during the season, tribes merge and the remaining survivors compete for individual immunity. Survivors voted off may be part of the jury. When there are a small number of remaining survivors who are finalists (typically two or three), the jury votes in favor of one of them to become the Sole Survivor who receives a cash prize of one million dollars. Figure 1 represents a graphical depiction of the voting history of a season of U.S. Survivor.
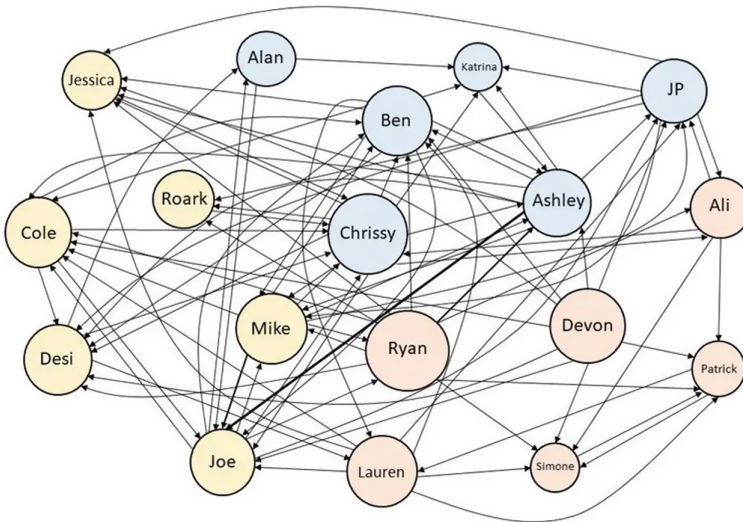
Fig. 1. The Survivor Heroes vs. Healers vs. Hustlers co-voting network. Nodes are scaled by closeness, and color-coded according to their original tribe. Thicker edges represent multiple votes.

Table 1. Three international Survivor seasons. Players are listed by first name, in order from top to bottom with the winner at the top, and the first eliminated player at the bottom. For each player we list the in-degree, out-degree, closeness, and CON score. The horizontal line separates finalists from the rest of the group.

Australian Survivor (2002)				Robinson 2009					Survivor South Africa Malaysia			
Name	ID	OD	C	CON	Name	ID	OD	C	CON	Name	ID	OD	C	CON

Australian Survivor (2002)					Robinson 2009					Survivor South Africa Malaysia				
Robert	5	10	0.714	44	Ellenor	0	6	0.563	36	Lorette	4	9	0.653	35
Sciona	1	9	0.652	37	Jarmo	7	8	0.557	34	Grant	5	8	0.653	33
Joel	7	8	0.625	35	Anna	2	10	0.645	54	Amanda	4	8	0.622	26
Katie	3	9	0.652	38	Nina	4	7	0.557	38	Mandla	0	8	0.652	32
Sophie	3	8	0.652	38	Erik Bl.	7	9	0.612	47	Angie	6	9	0.688	31
Jane	9	6	0.625	36	Lukas	4	7	0.557	31	Angela	4	6	0.594	28
Lance	8	5	0.577	27	Angela	6	8	0.612	46	Dyke	4	6	0.568	17
Craig	8	8	0.577	18	Ranjit	5	5	0.51	30	Hein	5	4	0.484	22
Naomi	8	7	0.5	25	Christian	3	4	0.51	24	Irshaad	3	5	0.544	25
Caren	10	6	0.5	25	Rafael	5	4	0.49	28	Lisa	11	4	0.484	16
Sylvan	3	5	0.417	30	Erik Bi.	9	5	0.438	26	Rijesh	4	3	0.408	13
Deborah	4	4	0.395	23	Erik R.	5	4	0.422	18	Nichal	6	2	0.363	12
Jeff	5	1	0.395	4	Mika	5	3	0.306	13	Elsie	8	2	0.436	9
David	6	3	0.441	23	Josefine	0	2	0.265	17	Viwe	5	2	0.344	11
Tim	4	2	0.294	10	Erika	7	2	0.306	15	Nicola	5	1	0.304	6
Lucinda	8	1	0.0	7	Beatrice	6	2	0.35	12	Nomfundo	4	1	0.335	8
					Micha	12	1	0.299	7					

We extend the analysis of the 35 U.S. seasons in [6] to 82 international seasons of Survivor. Data used in our analysis was obtained from the Survivor wiki pages https://survivor.fandom.com/wiki/Main_Page. Several seasons (beyond the 82) were excluded for varying reasons. In some cases, a wiki page exists, but there was no voting data. In other cases, much of the voting information was missing, or the rules are significantly different than the traditional version of the game shows. Nevertheless, the number of seasons collected exceeds the number in [6].

In Table 1 we display some of the CON scores for a few example seasons. We distinguish which players are finalists, since the rules change in determining who is the last player eliminated. For example, instead of eliminating the last player via votes *against* players, in survivor many players may return for a final vote *for* who they would like to win.

Table 2. Statistics on international Survivor seasons.

		CON	PageRank	Jaccard similarity	Random set
Survivor	**Top 3**	**57.3**	43.9	47.6	11.1–27.3
	Top 5	**81.7**	78.0	72.0	18.5–45.5

In Table 2, we detail relevant statistics on these networks. For each network, we consider whether the winner of the season had one of the top three or top five CON scores and list the percentage of such networks. For example from Table 2 we see that 81.7% of Survivor winners had one of the largest 5 CON scores. For comparison, we also compute PageRank (on the *reversed-edge* network, where we change the orientation of directed edge) and Jaccard similarity scores, which are both standard ranking scores. Jaccard similarity is a type of normalized CON score; see [10]. We find that the CON scores are a more accurate predictor for determining finalists of Survivor than both PageRank and Jaccard similarity. We observe that these results are consistent with the analysis in [6] for the U.S. seasons. As an added comparison, we list the probability of the winner appearing in a random set of three or five players; note that there is a range of percentages depending on how many players are in a given season. In the interest of space, we refer the reader to https://eikmeier.sites.grinnell.edu/uncategorized/ competition-show-data/ where we house all data on these seasons.
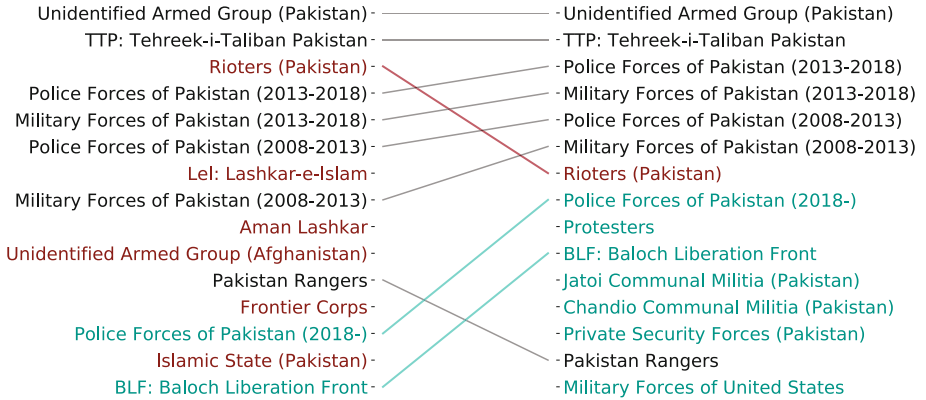
3.2   Political Conflicts

For our second competition network, we extracted data from *The Armed Conflict Location and Event Data Project* (or *ACLED*), which may be found at https:// www.acleddata.com/data/. ACLED catalogs information about political conflict and protests across the world. In these *conflict networks*, nodes correspond to actors in a given region, and edges correspond to conflicts between the actors. Many types of metadata are recorded corresponding to each event. Our particular interest is in the actors involved in each event, and where the event took place. More information about this project can be found on the project website.

An important note about the ACLED data is that we do not know which actor *initiated* a given event. Therefore, we do not consider the majority of edges (events) to be directed. The only events which we assume knowledge about directed-ness is when civilians are involved. We restricted our study to a set of events to a particular country; keeping the scale at the country level allows us to keep a larger set of actors. We selected three countries that have a large number of actors and events to analyze: Afghanistan, India, and Pakistan.

We first consider the rankings for Pakistan with commentary.

Pakistan has faced terrorism activities since 2000, with many militant groups attacking civilians and Pakistan armed forces. TTP (Pakistan) is one of the largest radical extremist groups, which is an umbrella organization of many militant groups such as Lashkar-e-Islam, Islamic State (Pakistan), and Jamaat-ul-Ahrar. In Fig. 2, we find that TTP has one of the highest CON scores. TTP has alliances with another terrorist organizations in Pakistan and neighboring countries, which lends to its prominence. In addition, due to the Afghan war, TTP has a strong influence and hold over many Islamic institutions in Pakistan. The Police Forces of Pakistan and Military Forces of Pakistan ensure national security, and they share information for achieving their goals. The Police Forces of Pakistan are an influential actor in the conflict network with another one of the highest CON scores. They perform their duties in all provinces of Pakistan

Unidentified Armed Group (Pakistan) — Unidentified Armed Group (Pakistan)
TTP: Tehreek-i-Taliban Pakistan — TTP: Tehreek-i-Taliban Pakistan
Rioters (Pakistan) — Police Forces of Pakistan (2013-2018)
Police Forces of Pakistan (2013-2018) — Military Forces of Pakistan (2013-2018)
Military Forces of Pakistan (2013-2018) — Police Forces of Pakistan (2008-2013)
Police Forces of Pakistan (2008-2013) — Military Forces of Pakistan (2008-2013)
Lel: Lashkar-e-Islam — Rioters (Pakistan)
Military Forces of Pakistan (2008-2013) — Police Forces of Pakistan (2018-)
Aman Lashkar — Protesters
Unidentified Armed Group (Afghanistan) — BLF: Baloch Liberation Front
Pakistan Rangers — Jatoi Communal Militia (Pakistan)
Frontier Corps — Chandio Communal Militia (Pakistan)
Police Forces of Pakistan (2018-) — Private Security Forces (Pakistan)
Islamic State (Pakistan) — Pakistan Rangers
BLF: Baloch Liberation Front — Military Forces of United States

Fig. 2. A Slope Graph to compare the rankings via CON and PageRank. On the left, the top actors in the conflict network Pakistan via CON metrics, while on the right, the top actors in Pakistan via PageRank on the reverse network. Actors are labeled in black if the difference in rankings is less than or equal to three. Actors are labeled in red if the CON ranking is at least four places higher than the PageRank, and in green if the PageRank is at least 4 places higher than the CON ranking. Note that no line appears to connect the left and right side if the actor does not show up in the top 15 of the other ranking.

with the help of their paramilitary forces such as Pakistan Rangers and Frontier Corps, and they maintain law and order, as well as border control. Military Forces of Pakistan (2013–2018) has one of the largest CON scores owing to their increased activities against terrorist groups in recent years.

We also offer commentary on some of the lower ranked actors. The Baloch Liberation Front (or BLF) is an ethnic-separatist political front and militant organization that is currently fighting against the Pakistani government for an independent Balochi state. The BLF is the strongest and most influential militant group of Baluchistan, but there has been no confirmed coordination between the BLF and other Balochi and non-Balochi groups, and they operate independently of one another. This is a large reason that BLF have low CON and closeness scores. The Islamic State is a part of the militant Islamist group: Islamic State of Iraq and Levant (ISIS). The Islamic State was formed by some of the TTP leaders and is more successful in Afghanistan. This organization has had less success in Pakistan largely carrying out isolated, small scale attacks. The Police Forces of Pakistan actively participated with the support of paramilitary forces of Pakistan in 2008–2018 for war against terror. The Police Forces of Pakistan mostly work to maintain the daily law and order in their respective provinces. Likely for these reasons, they have lower CON scores than the years between 2008–2018.
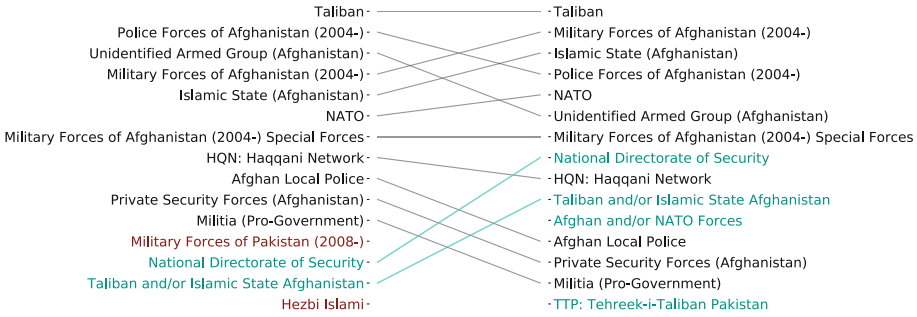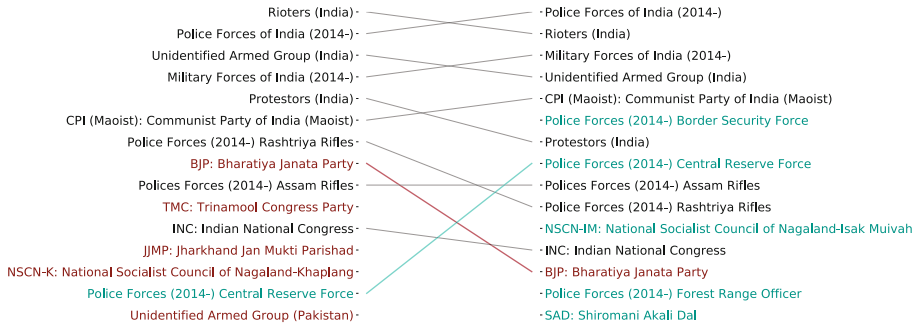
Fig. 3. Top actors in Afghanistan via CON score and PageRank.

We note that the ranking of the top actors using the CON score (on the left in Fig. 2) is not dissimilar to the one using PageRank on the reversed-edge network (on the right in Fig. 2). To quantify the difference in the rankings we used Spearman's rank correlation coefficient. Note that we cannot use Pearson correlation because our data is not at all Gaussian. Recall that Spearman's correlation coefficient is defined as

$$1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N(N^2 - 1)},$$

where N is the total number of actors, and d_i is the difference in rankings between actor i. A value close to 1 means that the two rankings are very well positively correlated. The Spearman correlation for Pakistan is -0.341, which suggests that the rankings are not that similar. In fact, the negative value implies that as the CON ranking decreases, the PageRank score increases. There are 741 total actors we consider in the Pakistan data set, and the later rankings clearly vary greatly.

Fig. 4. Top actors in India via CON score and PageRank.

We finish this section with the rankings for Afghanistan and India in Figs. 3 and 4. The Spearman coefficients are 0.604 and $-.267$ respectively, indicating

that the rankings provided by CON matches more similarly to PageRank in the Afghanistan dataset. While we do not provide in-depth commentary on these rankings, we find influential actors in both countries with the largest rankings against DCH metrics.

3.3   Food Webs

As a third and final type of data that we analyzed against the DCH, we studied food web datasets from the Pajek website: http://vlado.fmf.uni-lj.si/pub/networks/data/bio/foodweb/foodweb.htm [3]. These are 14 food webs in total. In food webs, the nodes are species, and a weighted edge (u, v) exists with weight w if u inherits carbon from v (that is, u preys on v) [2]. We interpret this as a directed negative interaction from node u to node v. A noteworthy difference in these networks (vs. Survivor, say) is that the movement of energy is *balanced*, meaning the in-degree and out-degree for each species is equal.
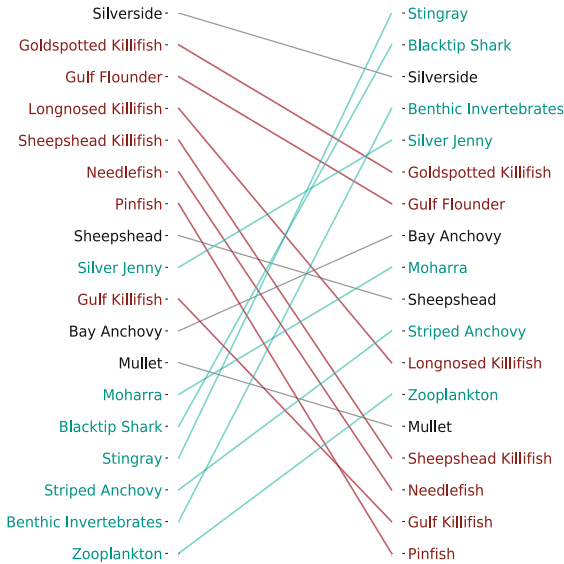

Fig. 5. The Chesapeake Bay Lower food web dataset. On the left, organisms are listed in decreasing order, with the largest CON score at the top. On the right, organisms are listed by decreasing PageRank score.

Rankings of selected food web datasets are in Figs. 5 and 6; the rankings for all the datasets may be found at https://eikmeier.sites.grinnell.edu/uncategorized/competition-show-data/ along with the computed CON scores, closeness, and PageRank on the reversed-edge network.

In studying the rankings of these 14 food webs, we see a difference between the CON rankings and PageRank. PageRank has been used to study the importance of species in regards to co-extinction [1,14,16], which we expect is likely

reflected in the rankings we see here using vanilla PageRank. However, we find a substantially different ranking when using the CON scores; for example, see the placement of Heteroflagellates in Fig. 5. The average Spearman correlation coefficient across these 14 datasets is 0.271, and the range is between 0.004 and 0.554. (Recall that a value close to 1 means very well correlated.) Therefore, we suggest that the CON scores are giving a *different* ranking, which is much closer to *trophic* levels of species. In particular, the CON scores reflect a natural hierarchical structure in ecosystems, and this is consistent with the DCH.

Fig. 6. The CrystalC food web dataset. On the left, organisms are listed in decreasing order, with the largest CON score at the top. On the right, organisms are listed by decreasing PageRank score.

4   Conclusion and Future Directions

We studied an implication of the Dynamic Competition Hypothesis (DCH) for competition networks across several different types of real-world networks. We found that the DCH prediction that high CON scores should correspond to leaders is supported in predicting winners in international seasons of Survivor, in predicting species with high trophic level species in food web, and for determining influential actors in conflict networks in Afghanistan, India, and Pakistan. Metrics such as CON scores outperformed PageRank as an indicator of influential actors in the competition networks we studied.

While our results provide support for the DCH, more work needs to be done. We did not address items (1) and (2) of the DCH regarding alliances in our data

sets, and that would be an important next step. Another direction is to consider an aggregate score, based on the CON score, closeness, and in- and out-degree, as a measure of detecting leaders in competition networks. An interesting direction would be to study more closely the dynamic aspects of competition networks, analyzing them over time to predict leaders. For example, we could analyze the co-voting network of Survivor of each episode of a season, and determine if temporal trends in network statistics predict finalists.

An open question is whether CON score centrality is applicable to large-scale networks exhibiting adversarial interactions, such as in Epinions and Slashdot (which give rise to signed data sets with tens of thousands of nodes, and available from [13]). Epinions was an on-line consumer review site, where users could trust or distrust each other. Slashdot is a social network that contains friend and foe links. A challenge with these data sets from the view of validating the DCH is that there is no inherently defined ranking, as there is in Survivor (via the order contestants were voted off), food webs (trophic level), and in conflict graphs (via political and strategic influence).

References

1. Allesina, S., Pascual, M.: Googling food webs: can an eigenvector measure species' importance for coextinctions? PLoS Comput. Biol. **59**, e1000494 (2009)
2. Baird, D., Ulanowicz, R.E.: The seasonal dynamics of the Chesapeake Bay ecosystem. Ecol. Monogr. **594**, 329–364 (1989)
3. Batagelj, V., Mrvar, A.: Pajek food web datasets. http://vlado.fmf.uni-lj.si/pub/networks/data/
4. Boginski, V., Butenko, S., Pardalos, P.M.: On structural properties of the market graph. In: Innovation in Financial and Economic Networks, Edward Elgar Publishers, pp. 29–45 (2003)
5. Bonato, A.: A Course on the Web Graph. American Mathematical Society Graduate Studies Series in Mathematics, Rhode Island (2008)
6. Bonato, A., Eikmeier, N., Gleich, D.F., Malik, R.: Dynamic competition networks: detecting alliances and leaders. In: Proceedings of Algorithms and Models for the Web Graph (WAW 2018) (2018)
7. Bonato, A., Tian, A.: Complex networks and social networks. In: Kranakis, E. (ed.) Social Networks. Mathematics in Industry Series. Springer, Berlin (2011)
8. Brandes, U., Erlebach, T. (eds.): Network Analysis: Methodological Foundations. LNCS 3418. Springer, Berlin (2005)
9. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets Reasoning about a Highly Connected World. Cambridge University Press, Cambridge (2010)
10. Gower, J.C., Warrens, M.J.: Similarity, dissimilarity, and distance, measures of, Wiley StatsRef: Statistics Reference Online (2006)
11. Guo, W., Lu, X., Donate, G.M., Johnson, S.: The spatial ecology of war and peace, Preprint (2019)

12. Heider, F.: The Psychology of Interpersonal Relations. John Wiley & Sons, Hoboken (1958)
13. Leskovec, J.: The Stanford large network dataset collection. http://snap.stanford.edu/data/index.html
14. McDonald-Madden, E., Sabbadin, R., Game, E.T., Baxter, P.W.J., Chadès, I., Possingham, H.P.: Using food-web theory to conserve ecosystems. Nat. Commun. **7**, 10245 (2016)
15. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Predicting positive and negative links in online social networks. In: Proceedings of the 19th International Conference on World Wide Web (WWW 2010) (2010)
16. Stouffer, D.B., Bascompte, J.: Compartmentalization increases food-web persistence. Proc. Nat. Acad. Sci. **1089**, 3648–3652 (2011)
17. Survivor Wiki. http://survivor.wikia.com/wiki/Main_Page
18. Tang, J., Chang, S., Aggarwal, C., Liu, H.: Negative link prediction in social media. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM 2015) (2015)
19. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)
20. West, D.B.: Introduction to Graph Theory, 2nd edn. Prentice Hall, New Jersey (2001)
21. Yang, S-H., Smola, A.J., Long, B., Zha, H., Chang, Y.: Friend or frenemy?: predicting signed ties in social networks. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2012) (2012)
22. Zachary, W.W.: An information flow model for conflict and fission in small groups. J. Anthropol. Res. **33**, 452–473 (1977)