PageRank as a Weak Tournament Solution*

Felix Brandt and Felix Fischer

Institut für Informatik, Universität München
80538 München, Germany
{brandtf,fischerf}@tcs.ifi.lmu.de

Abstract. We observe that ranking systems—a theoretical framework for web page ranking and collaborative filtering introduced by Altman and Tennenholtz—and tournament solutions—a well-studied area of social choice theory—are strongly related. This relationship permits a mutual transfer of axioms and solution concepts. As a first step, we formally analyze a tournament solution that is based on Google's PageRank algorithm and study its interrelationships with common tournament solutions. It turns out that the PageRank set is always contained in both the Schwartz set and the uncovered set, but may be disjoint from most other tournament solutions. While PageRank does not satisfy various standard properties from the tournament literature, it can be much more discriminatory than established tournament solutions.

1 Introduction

The central problem of the literature on tournament solutions is as appealing as it is simple: Given an irreflexive, asymmetric, and complete binary relation over a set, find the "maximal" elements of this set. As the standard notion of maximality is not well-defined in the presence of cycles, numerous alternative solution concepts have been devised and axiomatized [see, *e.g.*, 14, 12]. In social choice theory, the base relation, which we call dominance relation, is usually defined via pairwise majority voting, and many well-known tournament solutions yield attractive social choice correspondences. Recently, a number of concepts have been extended to the more general setting of *incomplete* dominance relations [9, 17, 6, 5]. These generalized dominance relations are commonly referred to as *weak tournaments*.

Motivated by the problem of ranking web pages based solely on the structure of the underlying link graph, Altman and Tennenholtz [3] introduced the notion of a *ranking system*, which maps each (strongly connected) directed graph to a complete preorder on the set of vertices. Obviously, this notion is strongly related to that of a tournament solution. In fact, Moulin [14] identifies "ranking the participants of a given tournament" as an important open problem. While little effort has been made so far to solve this problem, this is precisely what *ranking systems* achieve for strongly connected weak tournaments. Altman and

* This material is based upon work supported by the Deutsche Forschungsgemeinschaft under grant BR 2312/3-1.

Tennenholtz do not refer to the vast literature on tournament solutions, and their recent work on ranking systems does not seem to be well known in the tournament community. This is regrettable for two reasons. For one, ranking systems address a problem that has long been neglected in social choice theory. Secondly, both research areas could benefit from a mutual transfer of concepts and axioms. We take a first step in this direction by formally analyzing a tournament solution that is based on Google's PageRank ranking system.

2 Preliminaries

2.1 Weak Tournament Solutions

Fix an infinite set $\mathcal{A}$. A *weak tournament* is a pair $G = (A, \succ)$ of a finite set $A \subseteq \mathcal{A}$ of *alternatives* and an irreflexive and asymmetric *dominance relation* $\succ \subseteq A \times A$. Intuitively, $a \succ b$ means that a "beats" b in a pairwise comparison. We write $\mathcal{T}$ for the set of all weak tournaments, $\mathcal{T}(A)$ for the set of all weak tournaments on A, and $G|_{A'} = (A', A' \times A' \cap \succ)$ for the restriction of $G \in \mathcal{T}(A)$ to a subset $A' \subseteq A$ of the alternatives. A weak tournament is also called a *dominance graph*, and a weak tournament $(A, \succ)$ is a *tournament* if $\succ$ is complete. In the presence of (directed) cycles in the dominance relation, the concept of "best" or *maximal* elements is no longer well-defined, and various solution concepts that take over the role of maximality have been suggested. Some of these will be considered in Section 3. Formally, a *weak tournament solution* is a total function $S : \mathcal{T} \to 2^{\mathcal{A}} \setminus \{\emptyset\}$ such that for all $G \in \mathcal{T}(A)$, $S(G) \subseteq A$. We further require S to commute with any automorphism of $\mathcal{A}$, and to select the *maximum*, *i.e.*, an alternative that dominates any other alternative, whenever it exists.

2.2 The PageRank Set

PageRank assigns scores to pages on the Web based on the frequency with which they are visited by a "random surfer" [7, 15]. Pages are then ranked in accordance with these scores. It is straightforward to apply a similar idea to dominance graphs, starting at some alternative and then randomly moving to one of the alternatives that dominate the current one. Intuitively, this corresponds to a contestation process where the status quo is constantly being replaced by some dominating alternative. Arguably, alternatives that are chosen more frequently according to this process are more desirable than alternatives that are chosen less frequently.[1] A tournament solution based on PageRank should thus choose the alternatives visited most often by an infinite random walk on the dominance graph.[2]

[1] The key idea of this procedure is much older than PageRank and goes back to work by Daniels [8] and Moon and Pullman [13].

[2] It should be noted that transitions take place in the *reverse* direction of the dominance relation, from a dominated to a dominating alternative.

More formally, let $G = (A, \succ) \in \mathcal{T}(A)$ be a dominance graph, and let $d(a, G) = \{\, b \in A \mid a \succ b \,\}$ denote the dominion and $\overline{d}(a, G) = \{\, b \in A \mid b \succ a \,\}$ the dominators of alternative $a \in A$. Further let $\alpha \in [0, 1]$ be a parameter called the *damping factor*. Applying the original definition of Page et al. [15] to dominance graphs, the *PageRank score* $pr_\alpha(a, G)$ of alternative a in G is given by

$$pr_\alpha(a, G) = \alpha \left(\sum_{b \in d(a,G)} \frac{pr(b, G)}{|\overline{d}(b, G)|} \right) + \frac{(1 - \alpha)}{|A|} \quad .$$

That is, the score of a is determined by the scores of the alternatives it dominates, normalized by the number of alternatives dominating these, plus a constant.

It is well known that a solution to this system of equations such that $\sum_{a \in A} pr_\alpha(a, G) = 1$ corresponds to a stationary distribution of a Markov chain, and that a unique stationary distribution exists if the chain is irreducible, *i.e.*, if the dominance graph is strongly connected [see, *e.g.*, 11]. Undominated alternatives in the dominance graph lead to sinks in the Markov chain, thus making it irreducible. This problem can be handled by attaching either a self-loop or (uniform) transitions to all other states in the Markov chain to these sinks. The latter method, being the one commonly used in web page ranking, is clearly undesirable in the context of tournament solutions: For example, an undominated alternative that dominates some alternative inside a strongly connected subgraph would no longer be selected. Instead, we obtain the transition matrix of the Markov chain by transposing the adjacency matrix of the dominance graph, changing the diagonal entry to 1 in every row with sum 0, and row-normalizing the resulting matrix.

In the absence of sinks, pr_α is well-defined for every $\alpha < 1$. In the context of web page ranking, α has to be chosen carefully to accurately model the probability that a human user surfing the Web will stop following links and instead move to a random page [see, *e.g.*, 18]. Furthermore, the ability to differentiate between elements with lower scores is lost as α increases. The situation is different when PageRank is to be used as a tournament solution. In this case we want the solution to depend entirely on the dominance relation, and we are only interested in the best alternatives rather than a complete ranking. We thus want to compute pr_α for α as close to 1 as possible. It turns out that $\lim_{\alpha \to 1} pr_\alpha(a, G)$ is always well-defined [4], and we arrive at the following definition.

Definition 1. *Let $G \in \mathcal{T}(A)$ be a weak tournament. The* PageRank score *of an alternative $a \in A$ is defined as $pr(a, G) = \lim_{\alpha \to 1} pr_\alpha(a, G)$ where $\sum_{a \in A} pr_\alpha(a, G) = 1$. The* PageRank set *of G is given by $PR(G) = \{\, a \in A \mid pr(a, G) = \max_{b \in A} pr(b, G) \,\}$.*[3]

Boldi et al. [4] further observe that $\lim_{\alpha \to 1} pr_\alpha$ must equal one of the (possibly infinitely many) solutions of the system of equations for pr_1. This can be

[3] Another tournament solution based on random walks in tournaments, called the *Markov set*, is described by Laslier [12]. While their definitions are similar, there exists a tournament with five alternatives for which the two solutions are disjoint.

used to relate the PageRank set to a well-known tournament solution called the Schwartz set.

Definition 2. *Let $G \in \mathcal{T}(A)$ be a weak tournament. A set $X \subseteq A$ has the* Schwartz property *if no alternative in X is dominated by some alternative* not *in X. The* Schwartz set *$T(G)$ is then defined as the union of all sets with the Schwartz property that are minimal w.r.t. set inclusion.*

We further write $\overline{T}(G)$ for the set of weak tournaments induced by the minimal subsets of A with the Schwartz property.

It is well known from the theory of Markov chains that *every* solution of the system of equations for $\alpha = 1$ must satisfy $pr_1(a, G) = 0$ for all $a \notin T(G)$, and $pr_1(a, H)/pr_1(b, H) = pr_1(a, G)/pr_1(b, G)$ if $a, b \in A'$ for some $H = (A', \succ') \in \overline{T}(G)$ [see, *e.g.*, 11]. We thus have the following.

Fact 1. *Let $G \in \mathcal{T}(A)$ be a weak tournament. Then, for all $a \in A \setminus T(G)$, $pr(a, G) = 0$, and for all $H = (A', \succ') \in \overline{T}(G)$ and $a, b \in A'$, $pr(a, H)/pr(b, H) = pr(a, G)/pr(b, G)$.*

In particular, $PR(G)$ can be determined by directly computing pr_1 for the (strongly connected) graph $G|_{T(G)}$ if $|\overline{T}(G)| = 1$, a property that always holds in tournaments. If there is more than one minimal set with the Schwartz property, relative scores of alternatives in *different* elements of $\overline{T}(G)$ may very well depend on the dominance structure outside the Schwartz set, and it is not obvious that scores can be computed directly in this case.

3 Set-Theoretic Relationships

It follows directly from Fact 1 that the PageRank set is always contained in the Schwartz set. We will now investigate its relationship to various other tournament solutions considered in the literature [see, *e.g.*, 12, 6, 5]. In particular, we look at the (bidirectional) uncovered set and three other solutions that are always contained in the uncovered set.

Definition 3. *The* uncovered set *$UC(G)$ of a weak tournament $G \in \mathcal{T}(A)$ is given by*

$$UC(A) = \{\, x \in A \mid y \,C\, x \text{ for no } y \in A \,\} \enspace,$$

where $a \,C\, b$ if $a \succ b$ and for all $c \in A$, $c \succ a$ implies $c \succ b$ and $b \succ c$ implies $a \succ c$.

The Banks set *$B(G)$ of G contains the maximal element of each linear subrelation of the dominance relation induced by a set of alternatives that is itself maximal w.r.t. set inclusion,* i.e., *$a_1 \in B(G)$ if there exists a subset $A' \subseteq A$, $A' = \{a_1, \dots, a_k\}$ such that*

(i) $a_i \succ a_j$ if $1 \le i < j \le k$, and
(ii) there is no $b \in A$ such that $b \succ a_i$ for all $1 \le i \le k$.

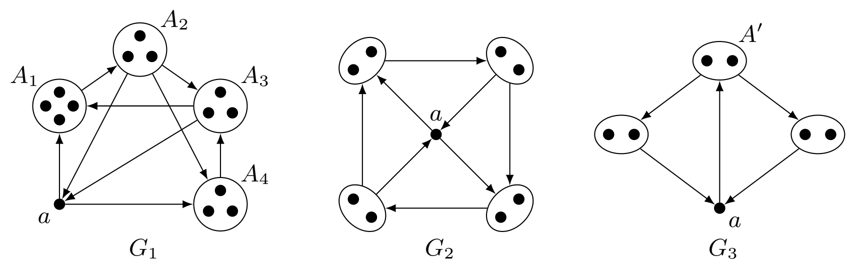

Fig. 1. $PR(G)$ can have an empty intersection with $UC_2(G)$, $B(G)$, $SL(G)$, and $C(G)$. An edge incident to a circle enclosing a set of alternatives is used to denote that there is an edge for each of these alternatives.

The Slater set $SL(G)$ *of* G *consists of the maximal elements of those acyclic relations that disagree with a minimal number of elements of the dominance relation,* i.e.,

$$SL(G) = \{ a \in A \mid \overline{d}(a, G') = \emptyset \textit{ for some } G' \in \operatorname*{argmin}_{G'' \in \mathcal{T}(A)} \Delta(G, G'') \} \quad ,$$

where $\Delta((A, \succ), (A, \succ')) = |\{ (a, b) \in A \times A \mid a \succ b \textit{ and } a \not\succ' b \}|$.

The Copeland set $C(G)$ *of* G *is the set of all alternatives for which the difference between the number of alternatives it dominates and the number of alternatives it is dominated by is maximal,* i.e.,

$$C(G) = \{ a \in A \mid a \in \operatorname*{argmax}_{a' \in A} |d(a', G)| - |\overline{d}(a', G)| \} \quad .$$

We further write $UC^k(G) = UC^{k-1}(G|_{UC(G)})$ for the kth iteration of the uncovered set. It is known that the Banks set intersects with all of these iterations, whereas $SL(G)$ may have an empty intersection with $UC^2(G)$. The main result of this section is stated next.

Theorem 1. *$PR(G)$ is always contained in $UC(G)$. $PR(G)$ may have an empty intersection with $UC^2(G)$, $B(G)$, $SL(G)$, and $C(G)$.*

Proof. For the inclusion in the uncovered set, we actually prove a stronger statement, namely that $PR(G) \subseteq UC_d(G)$, where $UC_d(G)$ is the set of downward uncovered elements of G, and $a \in A$ *downward covers* $b \in A$, denoted $a \, C_d \, b$, if $a \succ b$ and for all $c \in A$, $b \succ c$ implies $a \succ c$. Now consider $a \notin UC_d(G)$. By definition, there exists some $b \in A$ such that $b \, C_d \, a$. Furthermore, by Equation 2.2, $pr_\alpha(b) \geq pr_\alpha(a)$ for every $\alpha \in \,]0, 1]$, and thus $pr(b) \geq pr(a)$, with a strict inequality if $pr(a) > 0$. Since $pr(c)$ must be strictly positive for some $c \in A$, $a \notin PR(G)$ follows.

Now consider the dominance graphs in Figure 1. All three of them are strongly connected, such that the system of equations given by Equation 2.2 has a unique solution summing to 1 for $\alpha = 1$. It is straightforward but somewhat cumbersome to verify that the solution for G_1 is given by $pr_1(b, G_1) = p_b/1278$ where $p_a = 114$, $p_b = 93$ if $b \in A_1$, $p_b = 96$ if $b \in A_2$, $p_b = 112$ if $b \in A_3$, and $p_b = 56$ if $b \in A_4$. Thus, $PR(G_1) = \{a\}$. On the other hand, for all $b \in A_2$ and $c \in A_4$, b covers c in G_1, and for all $b \in A_3$, b covers a in $G|_{(A \setminus A_4)}$, such that $a \notin UC^2(G_1)$ (more precisely, $UC^2(G_1) = A_1 \cup A_2 \cup A_3$).

For G_2, the solution is given by $pr_1(a, G_2) = 4/28$, and $pr_1(b, G_2) = 3/28$ for all $b \in A \setminus \{a\}$. Thus, $PR(G_2) = \{a\}$. On the other hand, for every alternative $b \in A$ such that $a \succ b$, there exists an alternative $c \in A$ such that $c \succ a$ and $c \succ b$. Furthermore, there are no alternatives $b, c \in A$ such that $a \succ b$, $a \succ c$, and $b \succ c$, and thus $a \notin B(G_2)$ (more precisely, $B(G_2) = A \setminus \{a\}$).

Finally, the solution for G_3 is given by $pr_1(a, G_3) = 4/12$, $pr_1(b, G_3) = 2/12$ if $b \in A'$, and $pr_1(b, G_3) = 1/12$ if $b \in A \setminus (A' \cup \{a\})$, such that $PR(G_3) = \{a\}$. On the other hand, it is easily verified that $\{ (a, b) \mid b \in A' \}$ is the only set of two or fewer edges the removal of which makes G_3 cycle-free. At the same time, the members of A' maximize the difference between out- and in-degree over all vertices of G_3. Thus, $SL(G_3) = C(G_3) = A'$. □

4 Properties

In this section, we evaluate PageRank using standard properties from the literature on social choice theory and tournament solutions [see, *e.g.*, 9, 12]. Although some of the properties were originally introduced in the context of complete dominance relations, they naturally extend to the incomplete case.

In the following definition we use the notion of a component of a dominance graph. A nonempty subset $X \subseteq A$ of alternatives is called a component of $G \in \mathcal{T}(A)$ if for all $a, b \in X$, $d(a, G) \setminus X = d(b, G) \setminus X$ and $\overline{d}(a, G) \setminus X = \overline{d}(b, G) \setminus X$.

Definition 4. *A weak tournament solution S satisfies*

- monotonicity *if $a \in S(G)$ implies $a \in S(G')$ whenever $G|_{A \setminus \{a\}} = G'|_{A \setminus \{a\}}$, $d(a, G') \supseteq d(a, G)$, and $\overline{d}(a, G') \subseteq \overline{d}(a, G)$;*
- *the* strong superset property *(SSP) if for all $A' \subseteq A$, $S(G|_{A'}) = S(G)$ whenever $A' \supseteq S(G)$;*
- *the* Aïzerman property *if for all $A' \subseteq A$, $S(G|_{A'}) \subseteq S(G)$ whenever $A' \supseteq S(G)$;*
- idempotency *if $S(G|_{S(G)}) = S(G)$;*
- independence of the losers *if $S(G) = S(G')$ whenever $G|_{S(G)} = G'|_{S(G)}$;*
- weak composition-consistency *if for every $G \in \mathcal{T}(A)$, every component X of G and $a, b \in X$, and every $G' \in \mathcal{T}(A)$ such that $G|_{A \setminus \{a\}} = G'|_{A \setminus \{a\}}$ and $G|_{A \setminus \{b\}} = G'|_{A \setminus \{b\}}$, $S(G) \cap (A \setminus X) = S(G') \cap (A \setminus X)$ and $S(G) \cap X \neq \emptyset$ implies $S(G') \cap X \neq \emptyset$;*
- γ^* *if for all $A_1, \dots A_m \subseteq A$, $a \in S(G|_{A_i})$ for $1 \leq i \leq m$ implies $S(G|_{\cup_{i=1}^m A_i}) \neq \cup_{i=1}^m A_i \setminus \{a\}$.*

Property γ^* is a very weak version of (and implied by) the *expansion property* γ. Weak composition-consistency is a relaxation of composition-consistency, which in turn requires a tournament solution to select the best alternatives from the best components for any partition of the set of alternatives into components. Weak composition-consistence is weaker in that it does not require invariance of the choice under addition and removal of alternatives within a component. Finally, Aïzerman and idempotency are weakenings of SSP, their conjunction is equivalent to SSP.

Two properties of PageRank scores will be useful in the following. For the first property, consider a set $A' \subseteq A$ of alternatives with identical dominators and dominion, such that no member of the dominion is dominated by any alternative outside A'. If the size of A' increases, then the PageRank scores are distributed equally among all its members, while the PageRank scores of all other alternatives remain the same.[4]

Lemma 1. *Consider a dominance graph $G = (A, \succ)$ such that $|\overline{T}(G)| = 1$ and a set $A' \subseteq A$ of alternatives such that for all $a, b \in A'$, $\overline{d}(a, G) = \overline{d}(b, G)$ and $d(a, G) = d(b, G)$, and for all $b \in d(a, G)$, $\overline{d}(b, G) = A'$. Consider further a set A'' of alternatives with $A \cap A'' = \emptyset$ and the dominance graph $G' = (A \cup A'', \succ')$ such that for all $a, b \in A$, $a \succ' b$ if and only if $a \succ b$, and for all $a \in A'$, $b \in A''$, $\overline{d}(b, G') = \overline{d}(a, G)$ and $d(b, G') = d(a, G)$. Then, for all $a \in A \setminus A'$, $pr(a, G') = pr(a, G)$, and for all $a \in A' \cup A''$, $pr(a, G') = |A'|/(|A'| + |A''|)pr(a, G)$.*

Proof. Since $|\overline{T}(G)| = |\overline{T}(G')| = 1$, Equation 2.2 has a unique solution for $\alpha = 1$ in both cases. Denote $D = d(a, G)$ and $\overline{D} = \overline{d}(a, G)$ for some arbitrary $a \in A'$. The dominators and the dominion of each alternative $a \in A \setminus (A' \cup \overline{D})$ are the same in both G and G', such that $pr_1(a, G') = pr_1(a, G)$ if for all $b \in \overline{D}$, $pr_1(b, G') = pr_1(b, G)$. We show that the latter holds if for all $b \in D$, $pr(b, G') = pr(b, G)$, and prove the statement of the lemma in the process.

For all $a' \in A' \cup A''$ and $a \in A'$, we have

$$\begin{aligned} pr_1(a', G') &= \sum_{b \in d(a', G')} \frac{pr_1(b, G')}{|\overline{d}(b, G')|} = \sum_{b \in d(a', G')} \frac{pr_1(b, G')}{|A'| + |A''|} = \\ &= \frac{|A'|}{|A'| + |A''|} \sum_{b \in d(a', G')} \frac{pr_1(b, G')}{|A'|} = \\ &= \frac{|A'|}{|A'| + |A''|} \sum_{b \in d(a, G)} \frac{pr_1(b, G)}{|\overline{d}(b, G)|} = \frac{|A'|}{|A'| + |A''|} pr_1(a, G) \quad . \end{aligned}$$

[4] This bears some resemblance to the vote-by-committee axiom introduced by Altman and Tennenholtz [3].

Then, for all $a \in \overline{D}$ and $a' \in A'$

$$pr_1(a, G') = \sum_{b \in d(a,G')} \frac{pr_1(b, G')}{|\overline{d}(b, G')|} = |A'| \frac{pr_1(a', G')}{|\overline{D}|} =$$
$$= (|A'| + |A''|) \frac{|A'|}{|A'| + |A''|} \frac{pr_1(a', G')}{|\overline{D}|} = \sum_{b \in d(a,G)} \frac{pr_1(b, G)}{|\overline{D}|} =$$
$$= \sum_{b \in d(a,G)} \frac{pr_1(b, G)}{|\overline{d}(b, G)|} = pr_1(a, G) \quad .$$

□

It should be noted that the requirements of the lemma preclude any alternative $a \in A'$ from covering other alternatives (both upward and downward), because alternatives in $d(a)$ are not allowed to be dominated by alternatives in $\overline{d}(a)$, nor can a dominate an alternative in the dominion of a member of its own dominion. Hence, the lemma does not contradict the fact that covered alternatives cannot have maximum PageRank score.

The second property we are going to prove states that if a regular dominance structure is introduced for a set of alternatives with identical dominators and dominion, then the PageRank score of these alternatives increases, while the PageRank scores of all other alternatives decrease uniformly.

Lemma 2. *Consider a dominance graph $G = (A, \succ)$ with $|\overline{T}(G)| = 1$ and a set $A' \subseteq A$ of alternatives, $|A'| \geq 3$, such that for all $a, b \in A'$, $pr(a, G) > 0$, $d(a, G) = d(b, G)$, and $\overline{d}(a, G) = \overline{d}(b, G)$. Consider further the dominance graph $G' = (A, \succ')$ such that for all $a \in A$ and $b \in A \backslash A'$, $a \succ' b$ if and only if $a \succ b$ and $b \succ' a$ if and only if $b \succ a$, and for all $a, b \in A'$, $d(a, G'|_{A'}) = \overline{d}(b, G'|_{A'}) > 0$. Then, for all $a \in A'$, $pr(a, G') = (\overline{d}(a, G')/\overline{d}(a, G))pr(a, G)$, and for all $a, b \in A \setminus A'$, $pr(a, G')/pr(b, G') = pr(a, G)/pr(b, G)$.*

Proof. Again denote $\overline{D} = \overline{d}(a, G)$ and $D = \overline{d}(a, G)$ for some arbitrary $a \in A'$, and let $m = d(a, G'|_{A'})$. Consider a solution of the system of equations described by Equation 2.2 for G and $\alpha = 1$. Since G and G' only differ in the restriction to A', the same solution also satisfies Equation 2.2 for G' and all $a \in A \setminus (A' \cup \overline{D})$. Furthermore, by construction, for all $a \in A'$,

$$pr_1(a, G') = \sum_{b \in d(a,G')} \frac{pr_1(b, G')}{\overline{d}(b, G')} = \sum_{b \in d(a,G)} \frac{pr_1(b, G)}{\overline{d}(b, G)} + \frac{m}{m + \overline{D}} pr_1(a, G')$$

and thus

$$pr_1(a, G') = \frac{m + \overline{D}}{\overline{D}} pr_1(a, G) = (1 + \frac{m}{\overline{D}}) pr_1(a, G) \quad .$$

Then, for all $a \in \overline{D}$,

$$pr_1(a,G') = \sum_{b \in d(a,G')} \frac{pr_1(b,G')}{|\overline{d}(b,G')|} = \sum_{b \in A'} \frac{pr_1(b,G')}{|\overline{D}|} + \sum_{b \in d(a,G')\setminus A'} \frac{pr_1(b,G')}{|\overline{d}(b,G')|} =$$

$$= \sum_{b \in A'} \frac{\frac{m+\overline{D}}{\overline{D}} pr_1(b,G)}{m+|\overline{D}|} + \sum_{b \in d(a,G)\setminus A'} \frac{pr_1(b,G)}{|\overline{d}(b,G)|} =$$

$$= \sum_{b \in A'} \frac{pr_1(b,G)}{|\overline{D}|} + \sum_{b \in d(a,G)\setminus A'} \frac{pr_1(b,G)}{|\overline{d}(b,G)|} = pr_1(a,G) \quad .$$

Normalization of this solution yields the statement of the lemma. □

We use these properties to prove the main result of this section.

Theorem 2. *PR satisfies monotonicity. PR does* not *satisfy SSP, idempotency, Aïzerman, independence of the losers, weak composition-consistency, and* γ^**.*

Proof. As for monotonicity, it is obvious from Equation 2.2 that the score of an alternative $a \in A$ can never be decreased by having it dominate some alternative it did not dominate before, neither can this increase the score of some other alternative by more than the increase in the score of a. In particular, an alternative with maximum score will remain an alternative with maximum score.

For the Aïzerman property and idempotency, consider the dominance graph G obtained from a (directed) three-cycle by replacing one of the alternatives with a pair of alternatives, *i.e.*, $G = (\{a,b,c_1,c_2\},\succ)$ where $a \succ b$ and for $i \in \{1,2\}$, $b \succ c_i$ and $c_i \succ a$. By Lemma 1, $c_1 \notin PR(G)$. On the other hand, $G|_{\{a,b,c_1\}}$ is a three-cycle and thus $c_1 \in PR(G|_{\{a,b,c_1\}})$, violating the Aïzerman property. By the same argument, $PR(G) = \{a,b\}$ and thus $PR(G|_{PR(G)}) = \{a\}$, violating idempotency. Dutta and Laslier [9] show that SSP is equivalent to the conjunction of the Aïzerman property and idempotency, and the fact that PR does not satisfy the former follows from the fact that it does not satisfy the latter two.

For independence of the losers and weak composition-consistency, consider a set $A = \uplus_{i=1}^{3} A_i$ of alternatives, $|A_i| = 3$ for $i \in \{1,2,3\}$, and dominance graphs $G \in \mathcal{T}(A)$ such that each member of A_1 dominates each member of A_2, each member of A_2 dominates each member of A_3, and each member of A_3 dominates each member of A_1. If $G|_{A_1}$ is a three-cycle while $G|_{A_2}$ and $G|_{A_3}$ do not contain any edges, then, by Lemma 2, $PR(G) \cap (A_2 \cup A_3) = \emptyset$. This violates weak composition-consistency. If on the other hand $G|_{A_i}$ is a three-cycle for each set $i \in \{1,2,3\}$, then equation Equation 2.2 looks exactly the same for all alternatives, and $PR(G) = \cup_{i=1}^{3} A_i$. Together with the above, this violates independence of the losers.

For γ^*, consider the dominance graph G shown in Figure 2 and the sets $A_1 = \{a_1,a_2,a_3\}$ and $A_i = \{a_1,a_i\}$, $4 \le i \le 8$. It is easily verified that for $i = 1$ and for $4 \le i \le 8$, $a_1 \in PR(G|_{A_i})$. On the other hand, G is strongly connected and $pr_1(a_i,G)$ is defined for all i, $1 \le i \le 8$. Some computation

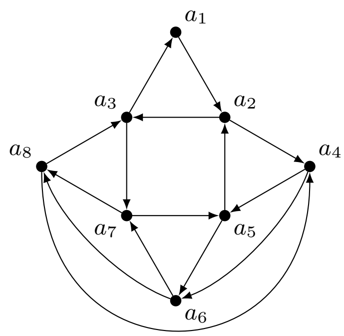

Fig. 2. PR does not satisfy γ^*.

yields $pr(a_1, G) = 1/15$ and for all i, $2 \leq i \leq 8$, $pr(a_i, G) = 2/15$. Thus, $PR(G) = A \setminus \{a_1\}$, violating γ^*. □

Let us now consider a property of tournament solutions that is in some sense orthogonal to the ones considered so far, namely *discriminatory power*. Indeed, the above properties describe which elements should be chosen given that some other elements are chosen as well, or should still be chosen as the overall set of alternatives changes. As we have shown, the PageRank set is uniformly smaller than both the Schwartz set and the uncovered set. We can again use Lemma 1 to establish that PageRank can be arbitrarily more discriminatory than every composition-consistent solution (*e.g.*, UC, UC^2, or B) in the sense that there exist instances where PageRank yields a singleton and any composition-consistent solution does not discriminate at all.

Theorem 3. *For any composition-consistent solution concept S and any set A of alternatives with $|A| \geq 5$, there exists a dominance graph $G \in \mathcal{T}(A)$ such that $|PR(G)| = 1$ and $S(G) = A$.*

Proof. Given a set A of alternatives, $|A| = k$, partition A into sets A_1, A_2, and A_3 with $|A_1| = 1$, $|A_2| = \lfloor (k - 1/2) \rfloor$, and $|A_3| = \lceil (k - 1/2) \rceil$ and let $G = (A, \succ)$ with $\succ = (A_1 \times A_2) \cup (A_2 \times A_3) \cup (A_3 \times A_1)$. Then, $S(G) = A$ due to the fact that S is invariant under automorphisms of $\{A_1, A_2, A_3\}$ and composition-consistent. On the other hand, A_2 and A_3 each contain at least two alternatives if $k \geq 5$, and, by Lemma 1, $PR(G) = A_1$. □

Similar properties, although less severe, can also be shown individually for solutions that are not composition-consistent, like the Slater set. We leave it as an open problem whether there exist dominance graphs in which PageRank yields a significantly larger choice than any of these sets.

It should finally be noted that PageRank has the advantage of being efficiently computable (if $|\overline{T}(G)| = 1$), whereas determining the Banks or the Slater set is NP-hard even in tournaments [20, 1].

5 Conclusion

The contribution of this paper is twofold. First, we identified a strong relationship between ranking systems and tournament solutions. Secondly, we formally analyzed PageRank using properties and solution concepts defined in the literature on tournament solutions. PageRank fails to satisfy a number of these properties, but on the other hand is very discriminatory—a well-known issue of most established tournament solutions [10]. It is open to debate whether these results cast doubt upon the significance of PageRank as a tournament solution, or the usefulness of some of the axiomatic properties used in the tournament literature.

An interesting problem for future work is to unify axioms in the literature on ranking systems [3, 2, 16, 19] and tournament solutions [*e.g.*, 12]. Some of these axioms are apparently based on very similar ideas.

Acknowledgements We thank Markus Holzer for helpful discussions and Paul Harrenstein for comments on a draft of this paper.

References

1. N. Alon. Ranking tournaments. *SIAM Journal of Discrete Mathematics*, 20 (1):137–142, 2006.
2. A. Altman and M. Tennenholtz. On the axiomatic foundations of ranking systems. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 917–922. Professional Book Center, 2005.
3. A. Altman and M. Tennenholtz. Ranking systems: The PageRank axioms. In *Proceedings of the 6th ACM Conference on Electronic Commerce (ACM-EC)*, pages 1–8. ACM Press, 2005.
4. P. Boldi, M. Santini, and S. Vigna. A deeper investigation of PageRank as a function of the damping factor. In *Web Information Retrieval and Linear Algebra Algorithms*, number 07071 in Dagstuhl Seminar Proceedings, 2007.
5. F. Brandt and F. Fischer. Computational aspects of covering in dominance graphs. In R. C. Holte and A. Howe, editors, *Proceedings of the 22nd Conference on Artificial Intelligence (AAAI)*, pages 694–699. AAAI Press, 2007.
6. F. Brandt, F. Fischer, and P. Harrenstein. The computational complexity of choice sets. In D. Samet, editor, *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 82–91. Presses Universitaires de Louvain, 2007.
7. S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 30(1-7):107–117, 1998.
8. H. E. Daniels. Round-robin tournament scores. *Biometrika*, 56(2):295–299, 1969.
9. B. Dutta and J.-F. Laslier. Comparison functions and choice correspondences. *Social Choice and Welfare*, 16(4):513–532, 1999.

10. M. Fey. Choosing form a large tournament. Mimeographed, University of Rochester, 2002.
11. J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Springer-Verlag, 1976.
12. J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer, 1997.
13. J. W. Moon and N. K. Pullman. On generalized tournament matrics. *SIAM Review*, 12(3):384–399, 1970.
14. H. Moulin. Choosing from a tournament. *Social Choice and Welfare*, 3: 271–291, 1986.
15. L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the Web. Technical Report 1999–66, Stanford University, 1999.
16. I. Palacios-Huerta and O. Volij. The measurement of intellectual influence. *Econometrica*, 72(3):963–977, 2004.
17. J. E. Peris and B. Subiza. Condorcet choice correspondences for weak tournaments. *Social Choice and Welfare*, 16(2):217–231, 1999.
18. L. Pretto. A theoretical analysis of Google's PageRank. In *Proceedings of the 9th International Symposium on String Processing and Information Retrieval (SPIRE)*, volume 2476 of *Lecture Notes in Computer Science (LNCS)*, pages 131–144. Springer-Verlag, 2002.
19. G. Slutzki and O. Volij. Scoring of web pages and tournaments—axiomatizations. *Social Choice and Welfare*, 26:75–92, 2006.
20. G. J. Woeginger. Banks winners in tournaments are difficult to recognize. *Social Choice and Welfare*, 20:523–528, 2003.