SIGMOD 2012
ACM SIGMOD 2012
May 20-24, 2012
Scottsdale, Arizona
USA

Reviews For Paper

Track	Research
Paper ID	201
Title	``Anything You Can Do, I Can Do Better." - Team Recommendation for Fantasy Games

---

Masked Reviewer ID: Assigned_Reviewer_1

Review:

Question	
Summary of the paper's main contributions and potential impact	The authors present the problem of fantasy team formation, wherein a team-creator has the option of choosing from a large number of players, all of whom have certain statistics. The goal is to choose a high-quality team (as measured by the team stats) from the player set. The authors model this as a skyline query, which they then attempt to answer using an apriori-style mechanism of removing dominated team candidates. I think the application goal here is a bit crazy and very far outside db interests, even liberally described. But the technical content is very interesting.
Strong points of the paper	I think the proposal of new non-monotonic properties is very interesting and could have very wide application. The authors view this as a core contribution of the paper and I agree. The experimental results on the NBA and synthetic datasets show good performance when using these properties. There's an interesting discussion in the experimental section. The content of the paper is very technical but generally easy to follow and well-written. The presentation of the paper is very good.

Weak points of the paper	The paper has one real weakness in my view, which is its limitation to the NBA dataset. This is a problem for two reasons: a) If team formation is the only application, then the impact of this paper is pretty limited. However, I think the order-specific property and the weak-candidate-generation property could have applicability to other non-team-oriented data problems. One possibility would be to choose tuples that compactly illustrate the contents of a much larger database. The authors should think more carefully about where this system could be applied elsewhere, and at least explain why some candidates are not worth considering. b) The power of the two properties is not really known until they are examined on other datasets. What if the # of attributes per object is very high? What if multiple attributes have values that are very close to each other? The robustness of the properties is not really examined. Thus this one decision I think limits the paper's potential impact, and the degree of confidence a reader has in the results.
Innovation	High
Quality of presentation	High
Quality and repeatability of experiments, if applicable	Satisfactory
Detailed comments	See above
Overall recommendation	Weak accept

Masked Reviewer ID: Assigned_Reviewer_2

Review:

Question	

Summary of the paper's main contributions and potential impact	The authors present techniques to identify groups of points which together dominate other groups of points. They applied these techniques to fantasy game team optimization.
Strong points of the paper	(S1) Interesting application. (S2) Technically quite interesting -- the identification of anti-monotonicity properties to speed up the algorithm. (S3) The paper is well-written with a good discussion on related work and the distinction from the current work.
Weak points of the paper	(W1) The formulation is a little weak -- the output size can be massive even larger than the number of points in the input. Hence, the output is unusable. (W2) The focus on scaling is perhaps not as important for the given application -- the number of players to choose from in these leagues is often only in thousands. Instead, the focus could have been on identifying aggregate functions where the output size is manageable for a human to look through.
Innovation	Medium
Quality of presentation	High
Quality and repeatability of experiments, if applicable	Satisfactory
Detailed comments	The application is fairly interesting. As mentioned above, it has certain characteristics that the size of the input is smaller. So, the emphasis on scaling with large sets of points is not important. Instead, the emphasis needs to be on understanding or further processing the output of these skyline formulations. In one of the experiments, the input size is 300 but the output size is already (for groups of size 3 each) 10000. So, one needs automated algorithms to further process the output. Expecting a human to look at these large output s and further choose the "best" or close to best is not feasible.
Overall recommendation	Weak reject

Masked Reviewer ID: Assigned_Reviewer_3

Review:

Question	
Summary of the paper's main contributions and potential impact	This paper addresses the problem of team recommendation for fantasy sports games. The authors model the problem as finding the skyline k-tuple groups from the given data. For example, game users may want the basket ball teams with 5 basket ball players out of all candidate players such that for each attributes such as score and rebound, the sum of the team members is at least that of any other possible teams with 5 players. Formally, the skyline k-tuple groups are the groups consisting of k tuples which are not dominated by any other groups of equal size. For the dominance relationship between two groups of k tuples, aggregate functions such as SUM, MAX and MIN can be used. Enumerating all possible k-tuples takes O(n^k) time. Thus, to reduce the search space, they provide two lemmas for anti-monotonicity properties and develop two algorithms to find skylines efficiently using their lemmas. With real-life datasets, the authors evaluate the performance of their algorithms with varying the number of the tuples in datasets and the number of the tuples per group. The results show that the proposed algorithms outperform the baseline method utilizing the general framework in [26]. To show the scalability of the proposed algorithms, the authors also experimented with synthetic datasets.
Strong points of the paper	S1: This is the first paper introducing the problem of team recommendations for fantasy games and the problem formulation to k-tuple skyline is very novel. S2: The authors present useful anti-monotonicity properties for frequently used aggregate functions with proofs. The properties are well developed and used well in their algorithms. S3: Presentation is good.
Weak points of the paper	W1: There is scalability problem with the proposed algorithms. Although the algorithms use pruning to reduce search space, the experimental results show that the number of the generated candidate groups increases rapidly as the number of tuples or the number of tuples per group grows. Because every generated candidate group is maintained in the main memory, these algorithms cannot handle large datasets.
Innovation	Medium
Quality of presentation	High

Quality and repeatability of experiments, if applicable	Satisfactory
Detailed comments	D1: As I pointed out in weak points, the authors should develop the proposed algorithms to handle the case when wither the number of tuples in data or the number of tuples per group is very large. D2: In 5th line of 2ndparagraph of page 2,, “pervious” should be “previous”. D1: Please show the result of execution times for synthetic data with increasing the number of tuples per group (i.e., k).
Overall recommendation	Weak accept