GameRank: Ranking and Analyzing Baseball Network

Zifei Shan, Shiyingxue Li, Yafei Dai
{shanzifei,lsyx09,dyf}@pku.edu.cn
北京大学
PEKING UNIVERSITY

Outline

- Introduction
- Algorithm and Evaluation
- Analysis and Visualization
- Future work and conclusion

INTRODUCTION

Background

- A baseball game:
  - two teams, take turns to attack and defend.
  - Players are batters in attacking phase, and pitchers/fielders in defending phase.
- Major League Baseball: the most attendance of any sports league. More than 70 million fans.
- Most previous research focuses on game video analysis.
- Full game records available on the Internet.

Questions

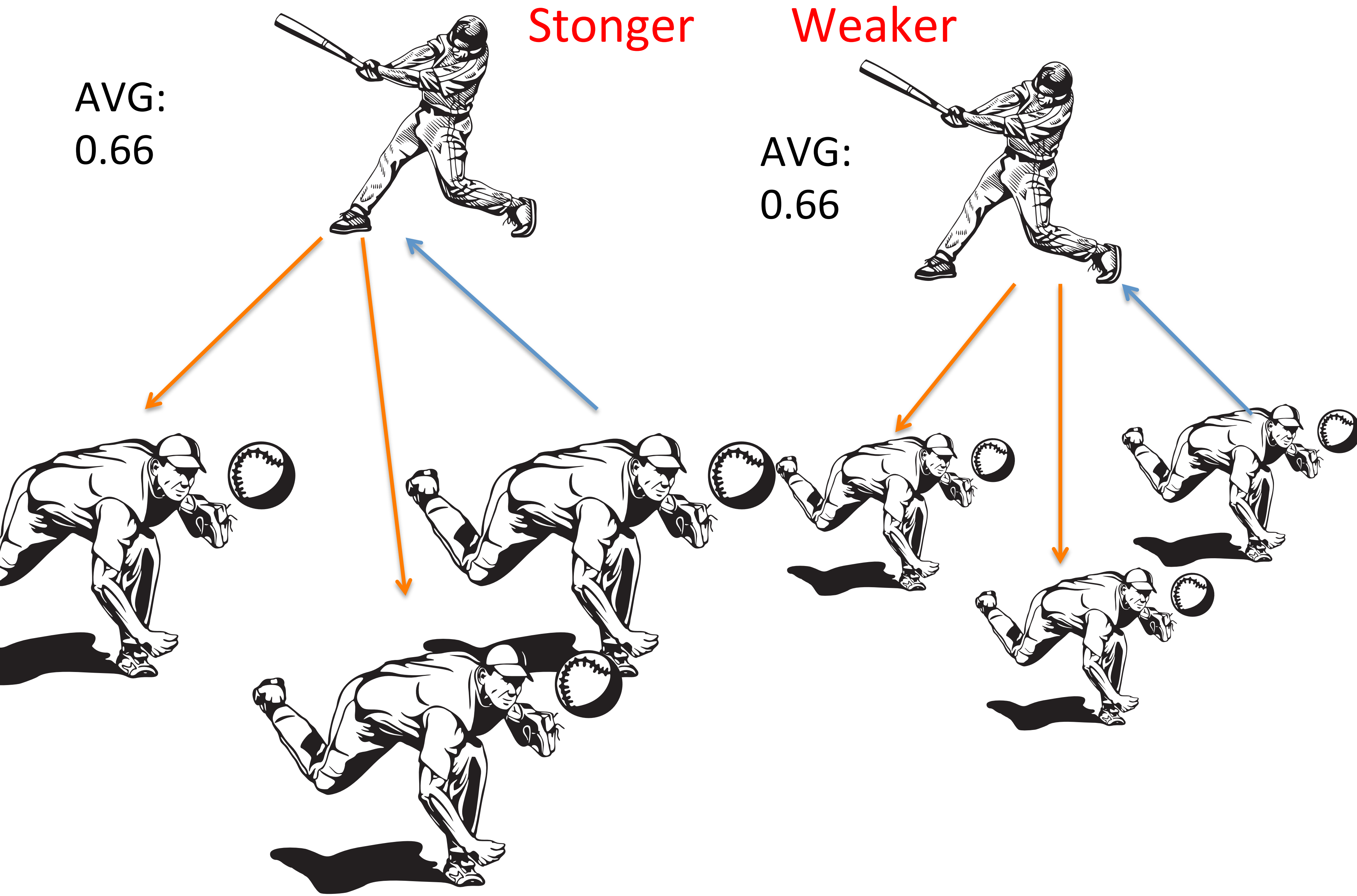
- How to rank baseball players?
- How to construct networks out of baseball games?
- What's special of baseball networks?
- What can we know from baseball network analysis?
- How about other sports networks?

Ranking Assumption

- Ranking players' pitching and batting ability separately:
  - a player is good at batting if he wins over good pitchers;
  - a player is good at pitching if he wins over good batters.
    - A good batter doesn't necessarily make (and usually isn't) a good pitcher.

Traditional Rankings

- Traditional Baseball Ranking:
  - Based on statistics
  - Hard to reflect the relationship of players.
  - E.g. Batting average:
    - Hits / at bats

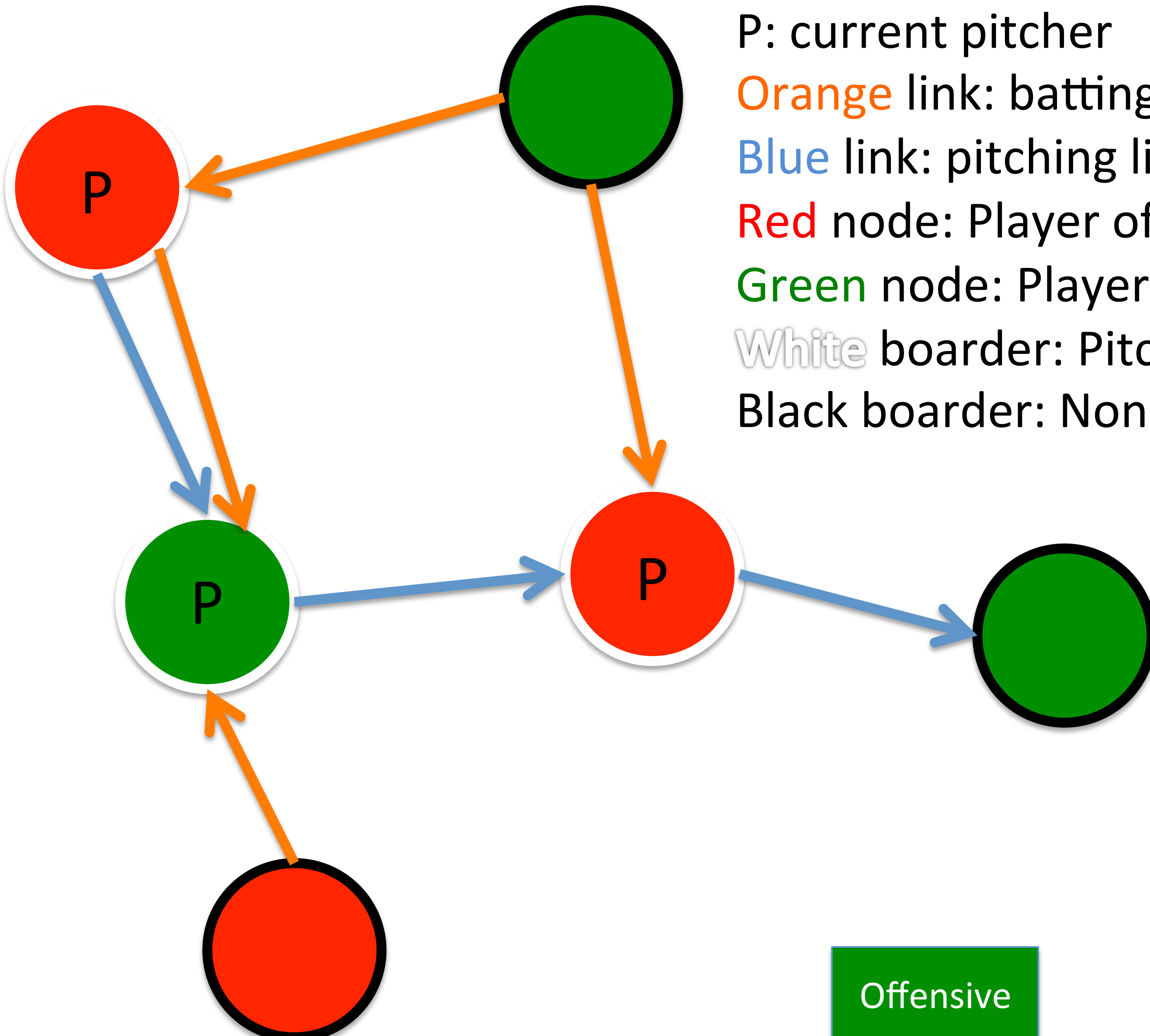

Stonger
Weaker
AVG:
0.66
AVG:
0.66

- So we want a model to take the relationships between players into consideration --- A network.

Network Construction

- Nodes ← Players
  - Two attributes: pitching ability, batting ability
  - A player can be a pitcher as well as a batter
- Links ← Win-lose relationships between players
  - Two types of links:
    - Pitching link A->B: A wins B when A is pitching
    - Batting link A->B: A wins B when A is batting

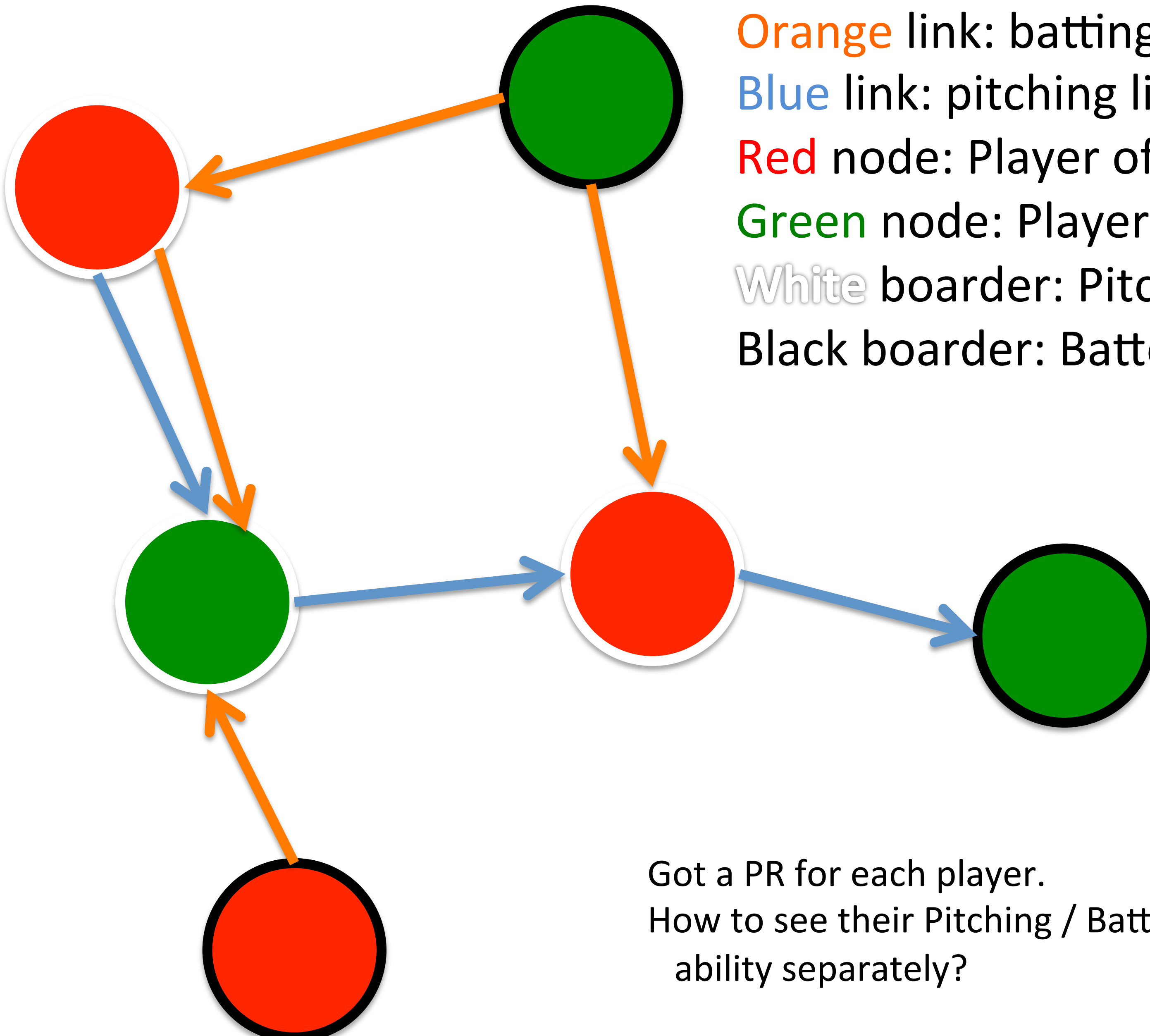
P: current pitcher
Orange link: batting link
Blue link: pitching link
Red node: Player of Team 1
Green node: Player of Team 2
White boarder: Pitcher
Black boarder: Non-pitcher



Player Ranking: PageRank?

- PageRank?
- Fail to separate two abilities: only have one indicator!
- See sample:

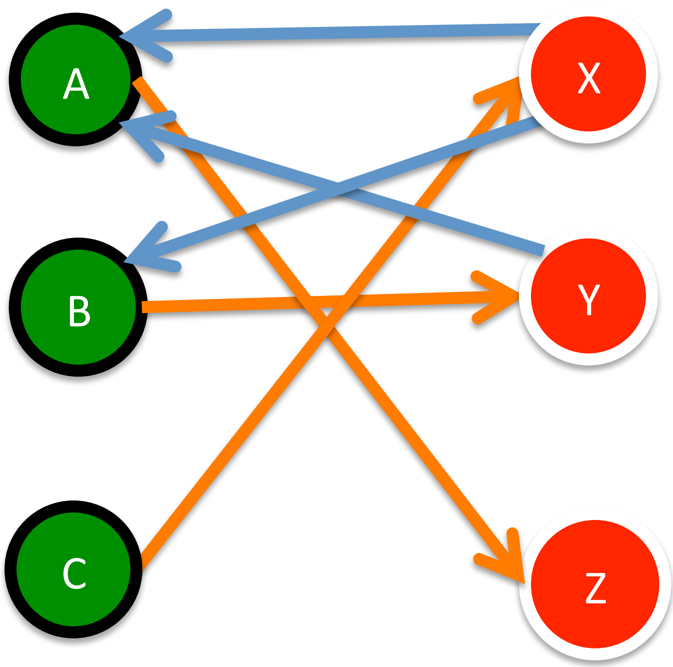
Orange link: batting link
Blue link: pitching link
Red node: Player of Team 1
Green node: Player of Team 2
White boarder: Pitcher
Black boarder: Batter



Got a PR for each player.
How to see their Pitching / Batting ability separately?

Player Ranking: Two PageRanks?

- Separate PageRank in two networks?
- Fail to describe the interplay between pitching and batting!
- See the following Sample:



Node size for green nodes: batting ability

Node size for red nodes: pitching ability

Orange link: batting link

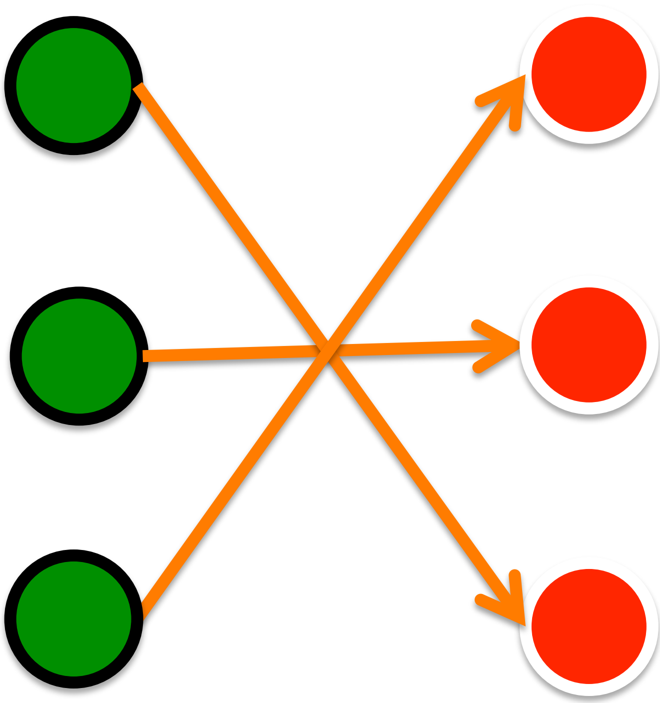
Blue link: pitching link

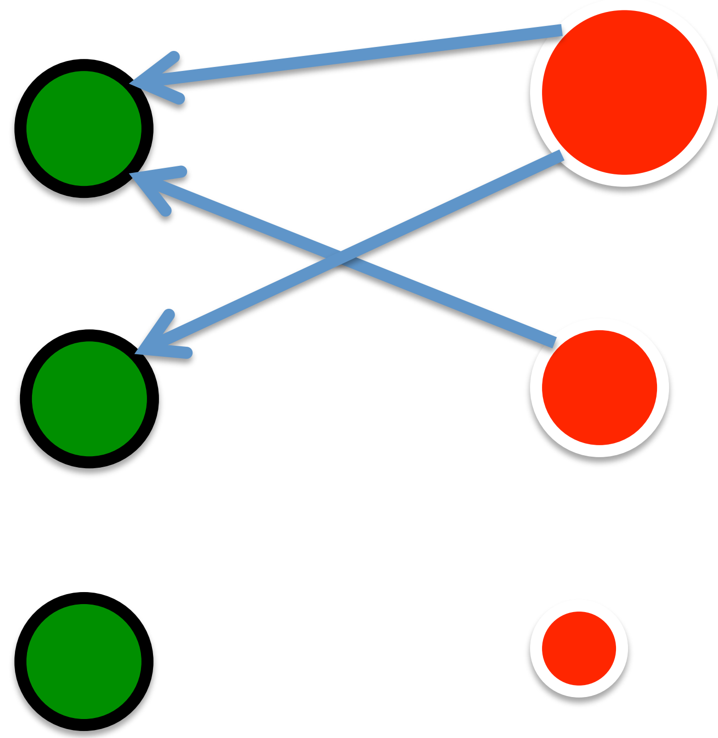
Red node: Player of Team 1 (all pitchers)

Green node: Player of Team 2 (all batters)

White boarder: Pitcher

Black boarder: Batter





Cannot distinct batters' abilities!

Player Ranking: HITS?

- We need a stronger ranking algorithm!
- HITS!
  - HITS: Hubs and authorities in Web
    - Good hubs links to good authorities
    - Good authorities are linked by good hubs
  - Similarly, baseball network:
    - Good pitchers wins good batters
    - Good batters wins good pitchers

Why not use HITS?

- We want two indicators that has **sound probabilistic meaning.**
- A random walk model like PageRank!

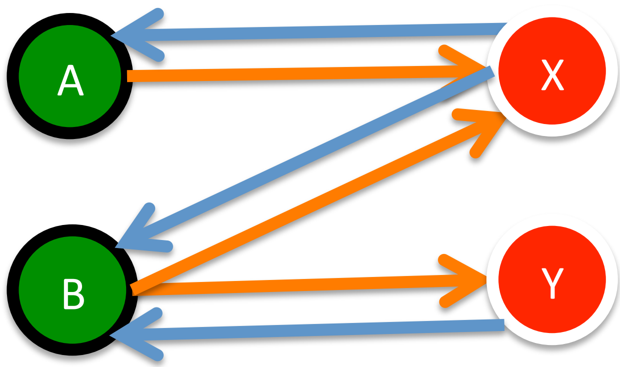
ALGORITHM: GAMERANK

GameRank: Overview

- We use the intuition of HITS, and build random walk models across the two (pitching and batting) networks.

Intuition: Random Walk

- Random walk in baseball (teams) network:
  - A baseball fan Ellie is trying to find the strongest player, by watching single plays through win-over relation (pitching/batting links) of players.
  - She starts randomly from batter A, and randomly picks a pitcher B who has won over A. And pick batter C who has won over pitcher B, etc.
  - If she finds a batter (pitcher) X that no one wins X, she will jump to a random pitcher (batter).
  - Sometimes she gets bored with the batter (pitcher) she's currently watching, and randomly picks another pitcher (batter).
- We can calculate The probability that she is watching a batter/pitcher after a long time = The frequency that she watches the player after a long time



GR	It1B	It1P	It2B	It2P	It3B	It3P
BA	1		0.5		0.375	
PA						
BB			0.5		0.625	
PB						
BX						
PX		1		0.75		0.687
BY						
PY				0.25		0.313

Orange link: batting link
Blue link: pitching link
Red node: Player of Team 1 (all pitchers)
Green node: Player of Team 2 (all batters)
White boarder: Pitcher
Black boarder: Batter

Definition

- Our formula:

$$GRB(i) = \beta/N - (1-\beta) \sum_{j \in outlinks_B(i)} \frac{GRP(j)}{DP_{in}(j)}, \qquad (1)$$

$$GRP(i) = \beta/N - (1-\beta) \sum_{j \in outlinks_P(i)} \frac{GRB(j)}{DB_{in}(j)}, \qquad (2)$$

  - $\beta = 0.15$

For Weighted Network

- Add edge weights
  - By modifying edge weights, we can make the rankings more precise with domain-specific knowledge

WEIGHT FOR DIFFERENT KINDS OF EDGES

Edge Class	Edge Type	Weight
Batting	Single Base	1
Batting	Double Base	2
Batting	Triple Base	3
Batting	Home Run	4
Batting	Sacrifice Hit	0.5
Batting	Walk / Base-on balls	0.5
Batting	Others	0.5
Pitching	All	1

Formula for weighted network

Then Batting Ability is

$$GRB(i) = \beta/N - (1-\beta) \sum_{j \in outlinks_B(i)} \frac{w_B(i,j)GRP(j)}{WDP_{in}(j)}, \tag{3}$$

Pitching Ability is

$$GRP(i) = \beta/N - (1-\beta) \sum_{j \in outlinks_P(i)} \frac{w_P(i,j)GRB(j)}{WDB_{in}(j)}, \tag{4}$$

Computation

- Start from a initial distribution, then iterately calculate GRB, GRP based on above formula.
- Will converge no matter what the initial distribution looks like.
- Can be easily parallelized with MapReduce model, similar to PageRank.

EVALUATION

Evaluation

- We evaluate our ranking algorithm in real-world, open-source MLB game records on *retrosheet.org*.
- We compare our result to ESPN Ratings, a prestigious ranking system.

Network of MLB data

- Pick year 2011 for evaluation
  - 1295 nodes
  - ~80000 aggregated edges
- Generate rankings for pitchers and batters with GameRank for 2011
- Get the ESPN ranks for 2011 from Internet

ESPN Ratings Algorithm

- ESPN Ratings uses a complex set of statistics.
  - E.g. the ESPN rating of batters includes the following factors: batting bases accumulated, runs produced, OBP, BA, HRs, RBIs, runs, hits, net steals, team win percentage, difficulty of defensive position, etc.
  - Hard to reflect relationships between players
- Not every player can get a ESPN score.

Comparison: Ranked Players

Ranking Algorithm	Ranked Batters	Ranked Pitchers
GameRank	823	659
ESPN	310	161

Comparison: top players

Top-10 Batters

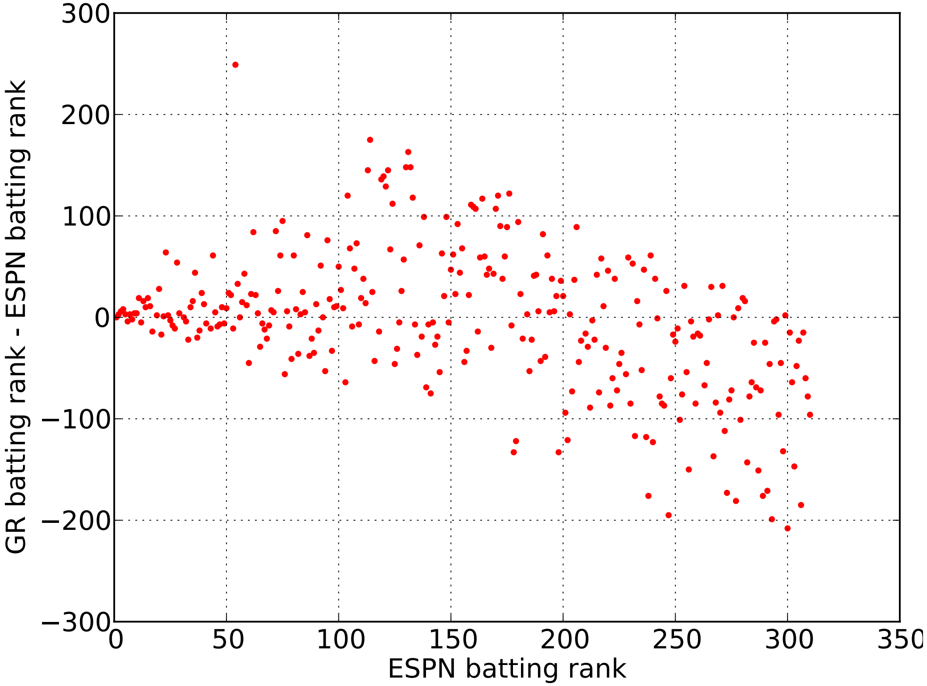
Name	GR Rank	ESPN Rank
Matt Kemp	1	1
Prince Fielder	2	6
Justin Upton	3	17
Hunter Pence	4	21
Ryan Braun	5	2
Joey Votto	6	8
Albert Pujols	7	12
Adrian Gonzalez	8	5
Jacoby Ellsbury	9	3
Jose Bautista	10	7

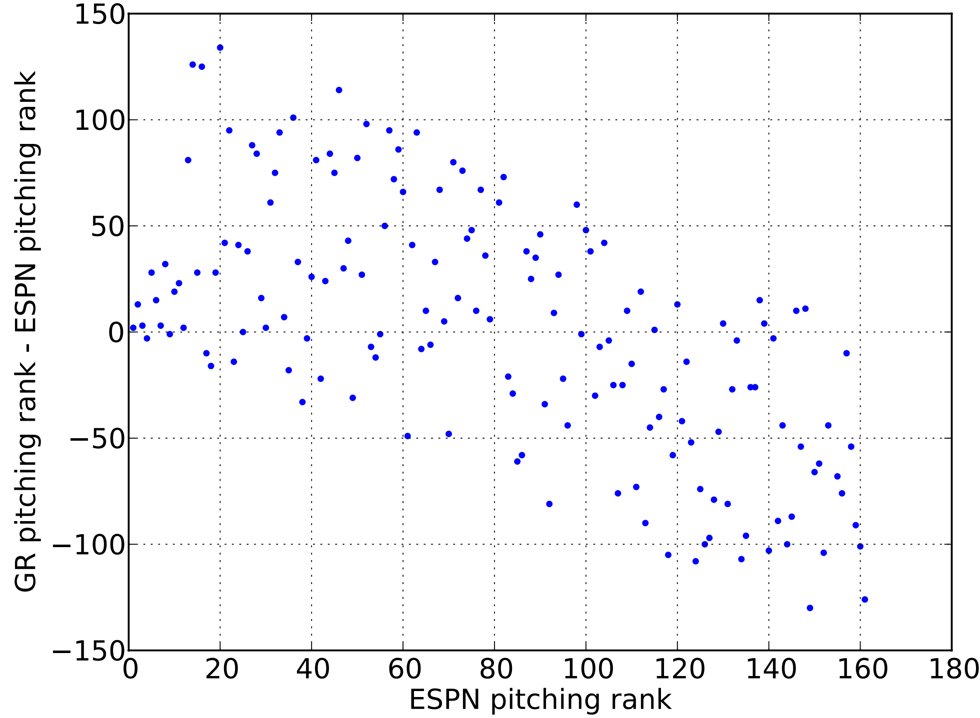
Top-10 Pitchers

Name	GR Rank	ESPN Rank
Cliff Lee	1	4
Matt Cain	2	18
Clayton Kershaw	3	1
Daniel Hudson	5	38
Roy Halladay	6	3
Tim Lincecum	7	17
Ian Kennedy	8	9
Tim Hudson	9	23
James Shields	10	7

- Top batters and pitchers found by GR, and their ESPN ranks.

Comparison: Difference



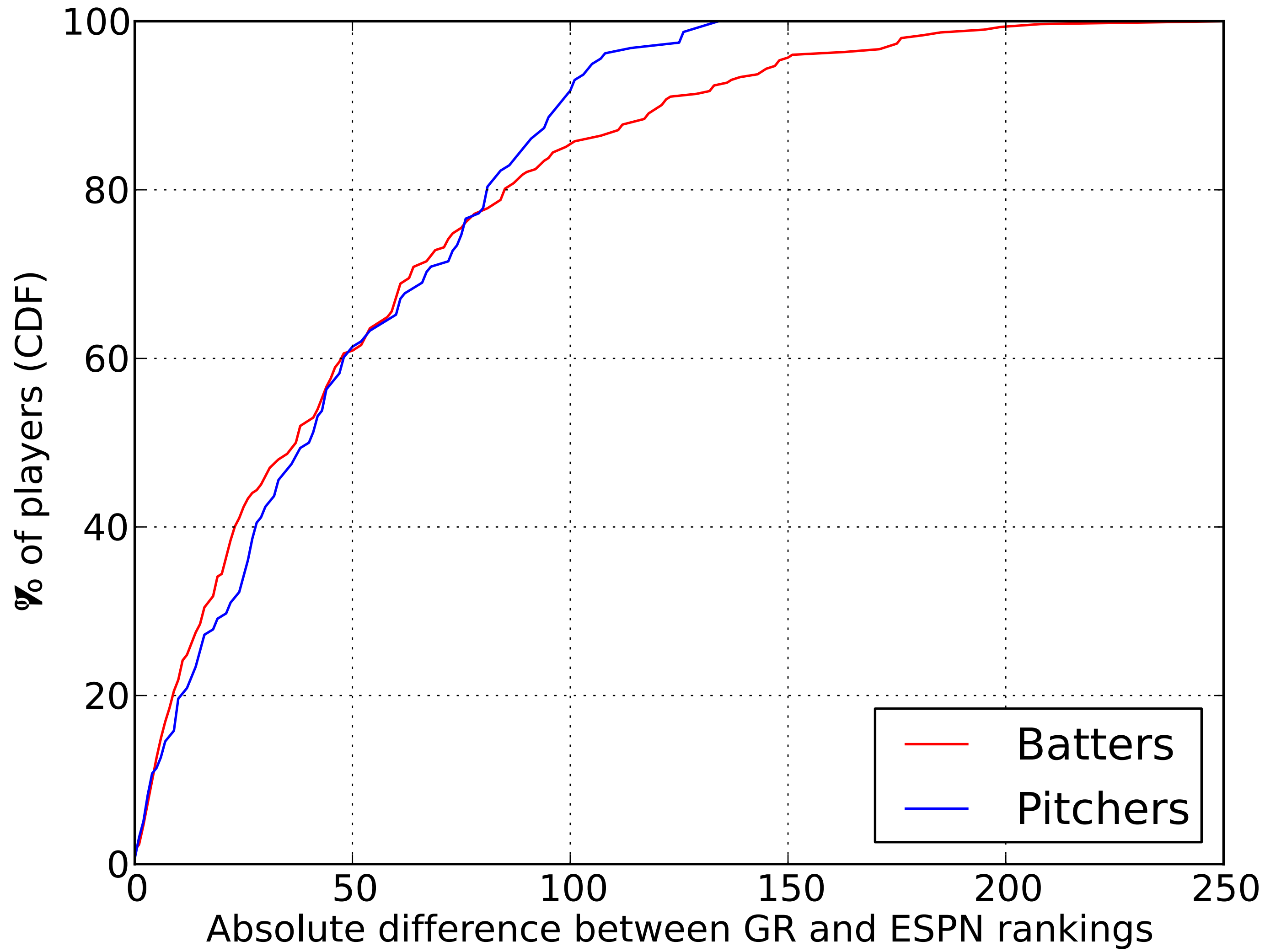


Batting

Pitching

(Scatter of difference between GR and ESPN)

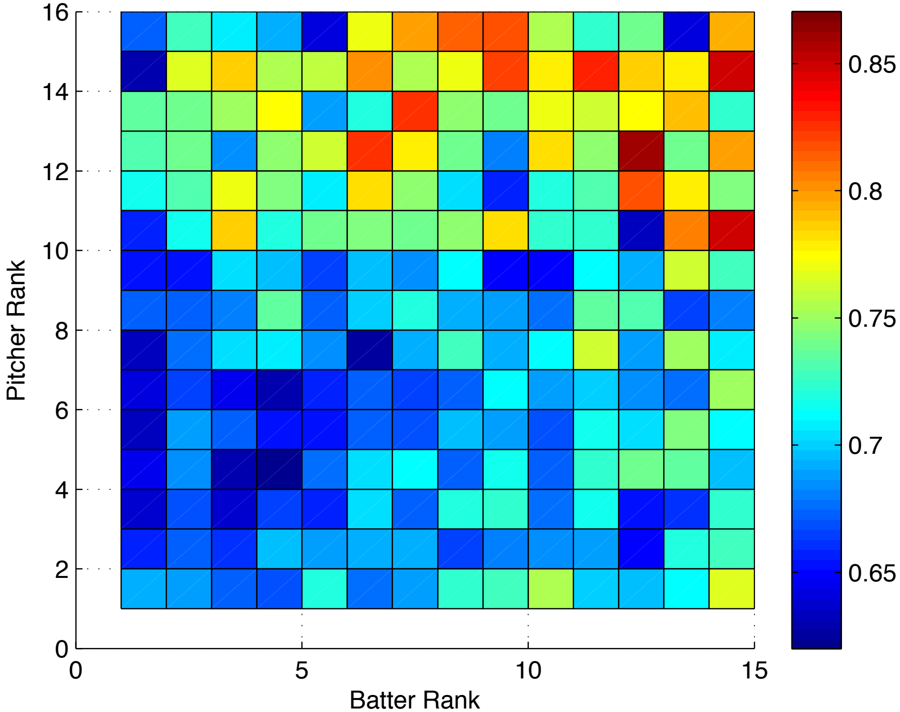
Comparison: Abs. Difference CDF



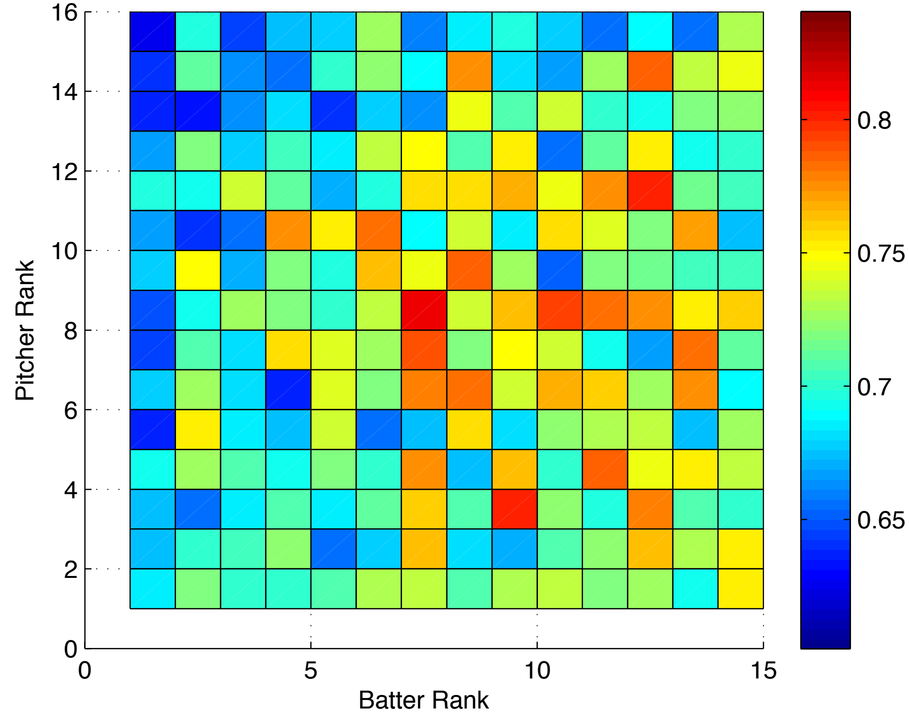
More comparison

- We already see that GR rankings achieves similar results with ESPN rankings.
- Now we want to prove that GR has better results than ESPN, with the intuition: players with better rankings should have higher probability to win in games.
  - if a ranking system is good, then under this system:
    - **Pitchers with high ranks** are more likely to win than **pitchers with low ranks**, and vise versa.
    - Pitchers at similar ranks are more likely to win **batters with low ranks than with high ranks.**

Comparison: Wining Rate



GR Rank



ESPN Rank

Frequency for pitchers to win batters at different rank levels in GameRank/ESPN. Pitcher ranks are divided by 10; batter ranks are divided by 20.

北京大學 PEKING UNIVERSITY

Evaluation: Conclusion

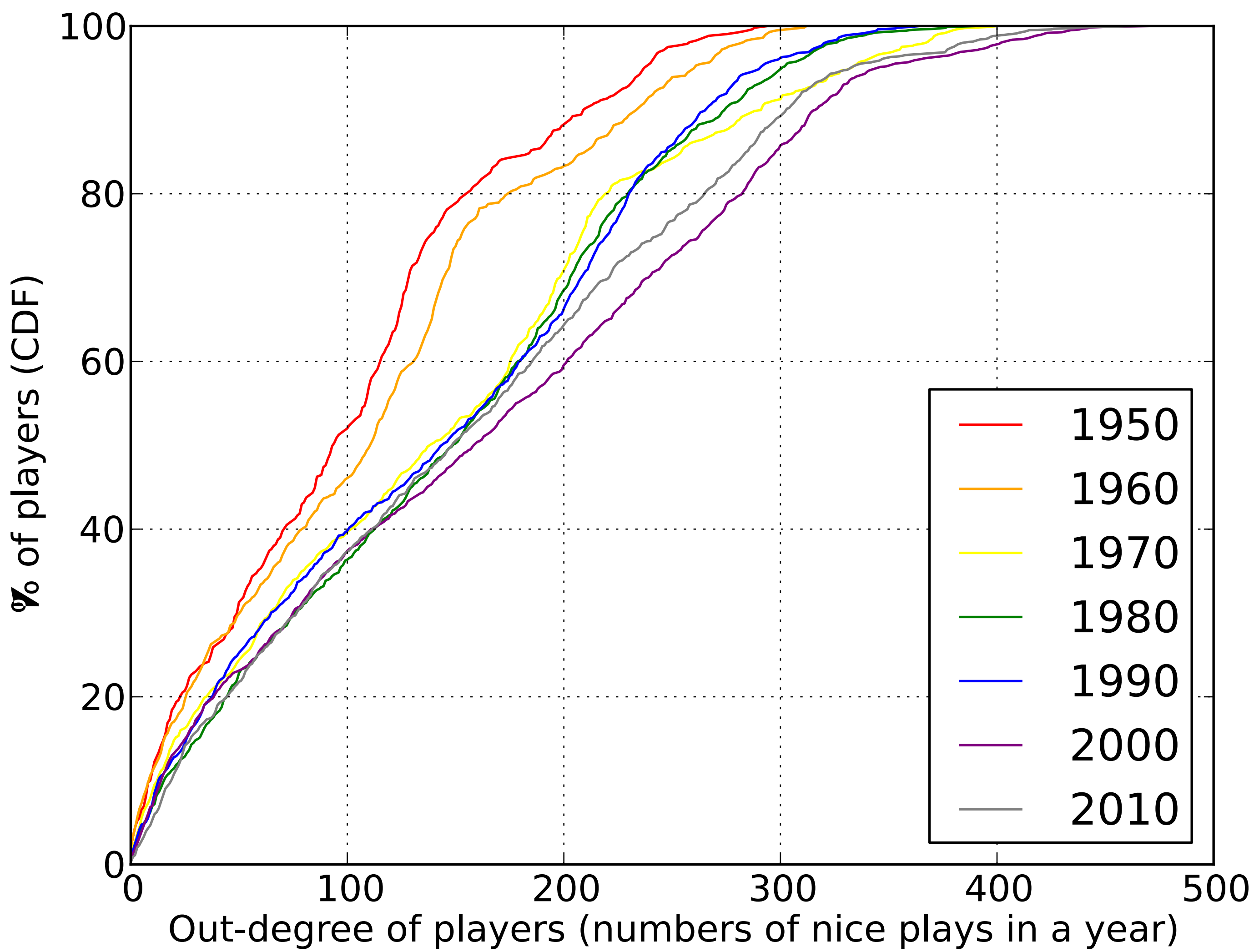
- GameRank achieves at least similar results with ESPN rankings
- GameRank is even better than ESPN in terms of batting rankings, if we set the criteria as wining frequency.
- GameRank can rank more (all) players.
- GameRank has a stronger model considering relationships between players.

ANALYSIS / DATA MINING

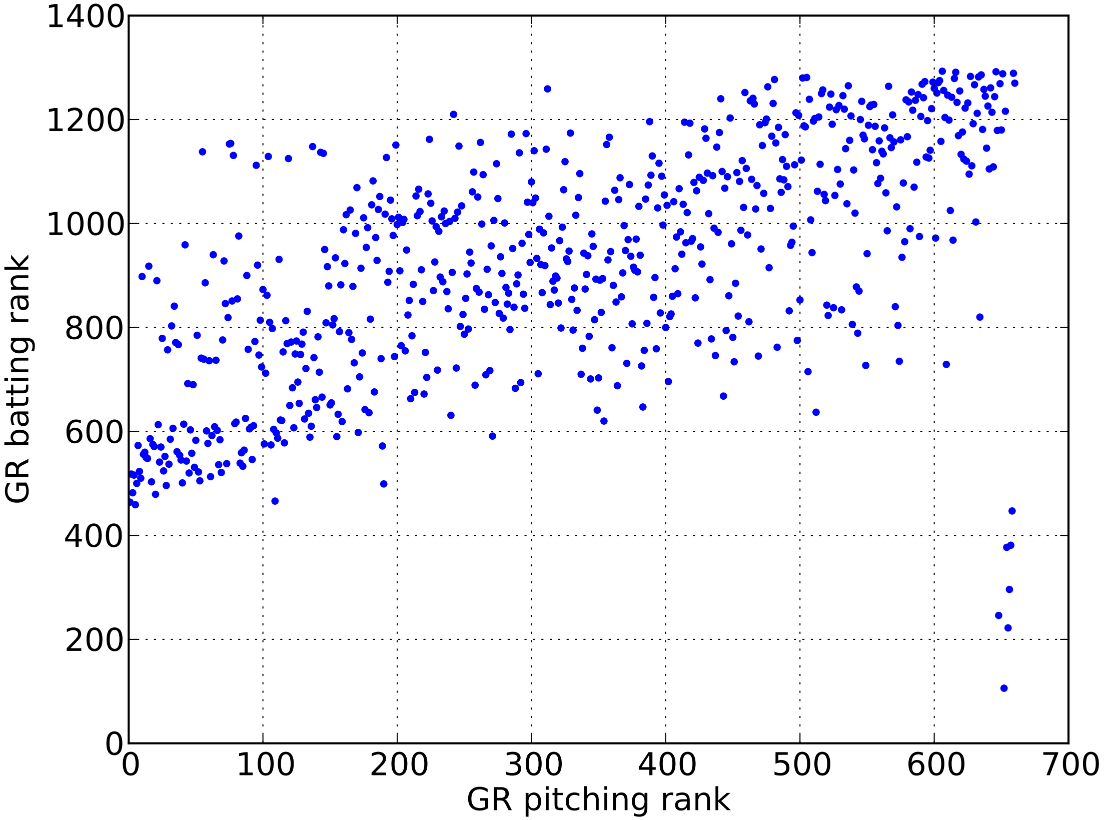
Analysis conclusions

- We analyze the networks with GR ranks, and found interesting results:
  – By studying the network's out-degree distribution in different years, we found that recent players are getting closer in their skills than before.
  – By analyzing the pitchers' GR batting values, we found that:
    - good pitchers are better than normal pitchers at batting.
    - Some bottom pitchers are great batters, because they do not usually pitch.

Analysis: out-degree distribution

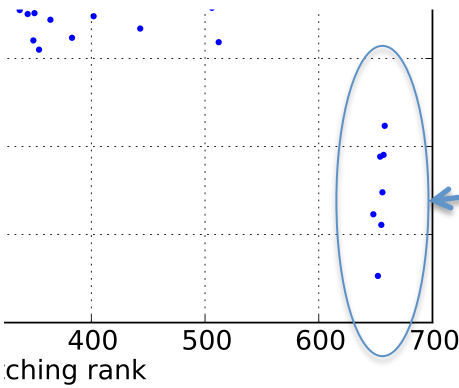


Analysis: Pitchers' batting ability



- Better pitchers bat better.

Analysis: bottom pitchers who bats well



BOTTOM PITCHERS WHO ARE GREAT BATTERS

Name	Batting Rank	Pitching Rank
Wilson Valdez	246	648
Michael Cuddyer	106	652
Darnell McDonald	377	654
Skip Schumaker	222	655
Bryan Petersen	296	656
Mike McCoy	381	657
Mitch Maier	447	658

- Among the bottom pitchers, there are 7 pitchers who bats really well.
  - We manually check them and found: most of them do not take pitchers as their **major fielding positions**, although they once pitched in 2011 regular season.

http://mlbillustrator.com

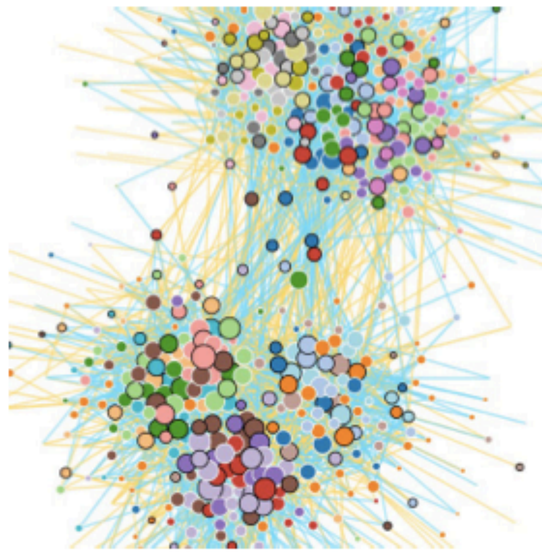
VISUALIZATION: MLBILLUSTRATOR

Visualization

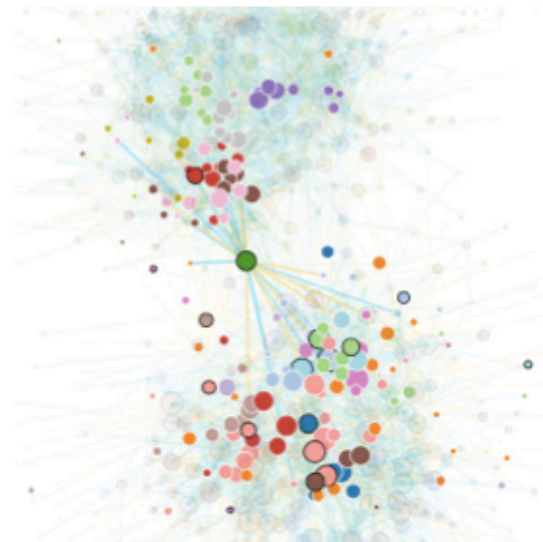
- We built an online website **MLBillustrator** to visualize the network and GameRank values for players:
  - http://mlbillustrator.com
- Then we do simple and initial analysis based on visualization.

Visualization

(a) Ranking Player by Team, 2009, Chicago White Sox with GameRank



(b) Ranking Player by ALL Teams, 2005 with GameRank



(c) A node and its neighbors in the network

Visual Analysis

- In every year, the network consists of two large communities.
  - Because in MLB there is American League (AL) and National League (NL), and the two clusters are almost exactly AL and NL communities.
    - Both AL and NL play more inside themselves, but less across leagues.
- Players in the middle of two communities: change teams across the league during the year.

OTHER USE CASES / FUTURE WORK / CONCLUSION

Other Use Cases

- GameRank algorithm is applicable for ranking networks with **multiple indicators interplaying with each other**.
- Other sports networks
  - Soccer
  - Volleyball
  - Basketball

Future work

- More analysis: find players that are overvalued/undervalued, etc.
- Test the robustness of each team in the network of in-team supports.
- Put players and teams into one heterogeneous network, and discover relationships between players and teams.
- Use specific knowledge in baseball games to optimize the parameters (edge weights).

Contribution

- We propose a ranking algorithm for networks with multiple indicators interplaying with each other.
- We initially regard baseball games as a network, and rank the pitching and batting ability of players.
- We analyze the baseball network and find interesting results.