

Computer system consistently makes most accurate NCAA picks

3 April 2008

Sports professionals and fans get pretty emotional about their picks for the NCAA basketball tournament each year, and that emotion often clouds their judgment.

But three engineering professors at the Georgia Institute of Technology have created a computer ranking system, called LRMC, that consistently predicts NCAA basketball rankings more accurately than the AP poll of sportswriters and the ESPN/USA Today poll of coaches, formulas (the Ratings Percentage Index), other computer models (the Massey ratings and the Sagarin ratings), and even the tournament seeds themselves.

After correctly picking all four of this year's finalists, the LRMC method has now identified 30 of the last 36 Final Four participants (83 percent accuracy over the past nine years of NCAA tournaments) as one of the top two teams in their region. Over the same nine-year stretch, the seedings and polls have correctly identified only 23, and the RPI identified 21.

LRMC (Logistic Regression Markov Chain) is a college basketball rankings system designed to use only basic scoreboard data, including which teams played, which team had home court advantage and the margin of victory. It was originally designed by Joel Sokol and Paul Kvam and has been maintained and improved by Sokol and George Nemhauser, all three optimization and statistics professors in the Stewart School of Industrial and Systems Engineering at Georgia Tech.

"As fans, we only get to see most tournament teams two or three times at most during the season, so our gut feelings about a team are really colored by how well or poorly they played the few times we've been watching," said Sokol. "On the other hand, our system objectively measures each team's performance in every game it plays, and mathematically balances all of those outcomes to

determine an overall ranking."

LRMC seems to have a particular knack for predicting good bubble teams and identifying the top teams. In addition to correctly picking the Final Four, LRMC also correctly identified several over-rated and under-rated teams as potential upsets. First-round losers Drake (5-seed, LRMC #30), Vanderbilt (4-seed, LRMC #38), and Connecticut (4-seed, LRMC #26), as well as second-round loser Georgetown (2-seed, LRMC #12), were all picked by LRMC as significantly over-rated teams.

On the other hand, teams like West Virginia (7-seed, LRMC #17), which defeated second-seeded Duke, and Kansas State (11-seed, LRMC #19), which defeated sixth-seeded USC, were correctly identified by LRMC as under-rated teams that could pull off one or more upsets.

But LRMC isn't perfect — it picked Clemson as under-rated (upset in the first round) and Davidson wasn't identified as under-rated by any major ranking method, including LRMC.

LRMC differs from other computer rankings systems in two important ways. When determining the value of home court advantage, LRMC considers how much playing at home helps a team win rather than how many points playing on a home court is worth.

Georgia Tech researchers have also been able to show that very close games are often "toss-ups," meaning the better team barely wins more than half the time. So, they determined that winning a close game shouldn't be worth as much as winning easily, and losing a close game shouldn't hurt a team's ranking as much as losing badly. LRMC's ranking methodology takes this into account.

Similar to other rankings systems, LRMC also uses the quality of each team's results and the strength of each team's schedule to rank teams.

So which team does LRMC favor for the top spot this year? It's chosen Kansas, despite UNC, UCLA and Memphis being the top three ranked teams by most systems.

Source: Georgia Institute of Technology

APA citation: Computer system consistently makes most accurate NCAA picks (2008, April 3) retrieved 22 October 2019 from <https://phys.org/news/2008-04-accurate-ncaa.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.