中国科学院　随机复杂结构与数据科学重点实验室　学术报告

题　　目: Foundational questions about sports rating models

报 告 人:　Prof. David Aldous

University of California Berkeley

时　　间：　2015 年 05 月 15 日（星期五）　下午 4:30-5:30

地　　点：　数学院南楼 902

摘　　要：　There is a natural toy probability model for sports results: each player [or team] has a “strength”, and the probability A beats B is a specified function W of their difference in strengths. This model underlies two well-established fields: more realistic modeling of professional sports, and the Bradley-Terry scheme for deriving a consensus ranking from inconsistent partial rankings. But one can imagine many other uses of the model. For instance

(i) does the standard design of a single-elimination tournament in terms of the seeding of players optimize anything?

(ii) In a given year, what is the probability that the winner of the Premier League or the Superbowl was actually the best team?

Our main focus is on Elo-type dynamical rating schemes, in which the rating of the winner [loser] is increased [decreased] by an amount which is a function Υ of the difference in ratings. Such schemes, very easy to implement, are used extensively for games with many amateur players and unsystematic scheduling of matches, for instance online games.

Such algorithms have nothing to do with probability, a priori, but there is an obvious heuristic connection with the toy probability model which (curiously) seems not to have been studied carefully. How accurate are such schemes if player strengths are unchanging, and how well do they track changes? Can one make more accurate schemes using two-component ratings, as implemented for instance in the TrueSkill ranking system on Xbox Live?