

<http://www.galaktika.ru/amm/introduction-to-scm.html> (дата обращения 04.04.2016).

3. Построение алгоритма отбора факторов, влияющих на конъюнктуру товарного рынка [Электронный ресурс] – Режим доступа. – URL <http://dis.ru/library/531/26996/> (дата обращения 04.04.2016).

ВЕРОЯТНОСТНАЯ МОДЕЛЬ РЕЙТИНГОВ ИГРОКОВ В КОМАНДНЫХ ВИДАХ СПОРТА

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Михолап А. А., Стержанов М. В.

Рассмотрена баесова модель оценки уровня игры спортивных команд и отдельных участников. Приведен способ уменьшения вычислительной сложности модели с помощью фактор-графа. Предложен способ повышения предсказательной точности путем агрегации результатов различных систем.

Ключевые слова: системы рейтингов, баесовы методы, вероятностные модели.

Задача ранжирования и построения рейтингов возникает в различных ситуациях и особенно часто в контексте спортивных соревнований. Объективная оценка навыков игроков и команд помогает принимать решения при покупке новых игроков, построении списка лучших команд и прогнозировании исходов будущих матчей.

В 2006 году корпорация Microsoft опубликовала статью с описанием новой рейтинговой системы TrueSkill. Она была разработана коллективом авторов для подбора соперников в многопользовательской онлайн игре Halo 2 и, согласно мнению авторов, может рассматриваться как обобщение Elo.

Недостаток Elo заключается в дискретных уровнях коэффициента скорости изменения рейтингов K . Они были подобраны вручную, на основании истории игр и здравого смысла. Однако, желательно, чтобы система могла подстраивать K автоматически. Например, если новичок играет со старым игроком, уровень игры которого стабилен и достаточно точно известен, то результат этого матча намного полезнее, чем матча против другого новичка, о котором ничего не известно. Чтобы сделать это очевидным, можно упростить и принять уровень навыка опытных игроков за точку. Тогда, сравнивая новичка с ними, мы получим для него довольно точные координаты. Если же случайная величина рейтинга противника также имеет широкий интервал, мы не можем сделать таких выводов, так как оценка очень неопределенная, а сами интервалы даже могут перекрываться.

TrueSkill помимо среднего значения μ варьирует также и дисперсию σ для каждого игрока, благодаря чему может автоматически сжимать и увеличивать интервал неопределенности без ручного вмешательства, основываясь лишь на истории игр.

Второй предпосылкой создания нового алгоритма была необходимость моделирования ничьих в явном виде. Вероятность ничьей это один из критериев качества разбиения игроков на команды в Halo 2 и других играх и видах спорта. Матч наиболее интересен, когда силы обеих сторон равны и исход непредсказуем.

Также в систему встроена детализация на уровне отдельных игроков. Чтобы автоматизировать управление случайными величинами приходится неизбежно усложнять алгоритм. Если в Elo использовалась базовая теория вероятности, логистическая функция и простое правило обновления, в TrueSkill за счет введения переменной дисперсии применяются значительно более сложные конструкции.

С начала двухтысячных годов в статистике начали набирать значительную популярность Баесовы методы. Они представляют собой набор правил работы со случайными величинами, позволяя упростить идеальную вероятностную модель до практически применимого уровня. Это возможно за счет ручного определения значимых факторов (которые описываются случайными величинами) и связей между ними.

TrueSkill использует именно Баесов подход, выражая модель встречи двух команд фактор графом (рисунок 1).

Система инициализируется определенными значениями, называемыми априорной информацией. В нее можно включить всевозможные сведения от экспертов, но в данном случае это просто фиксированный стартовый рейтинг. Важно различать априорную информацию в контексте всей системы и одного матча.

До игры осуществляется прямой проход по фактор графу сверху (от узлов, содержащих рейтинг игроков) вниз (к узлам, обозначающим исход матча). С его помощью оцениваются вероятности различных исходов. После завершения игры случайные величины в нижних узлах заменяются на реальные числовые

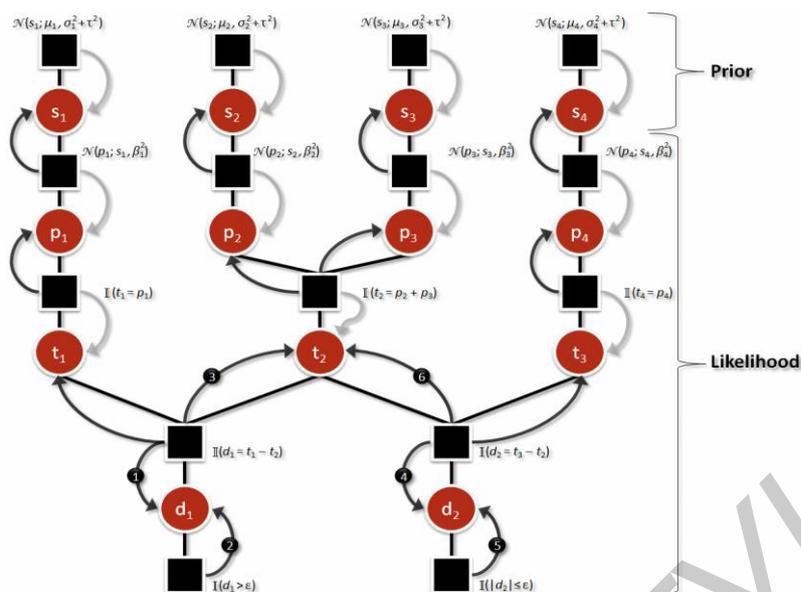


Рис. 1. Фактор граф TrueSkill

результаты и происходит обратный проход снизу вверх. Он обновляет рейтинги так, чтобы только что произошедшее событие стало более вероятным. То есть, если команда выигрывает — увеличивает рейтинг игроков, если проигрывает уменьшает.

Баесовы методы известны своей вычислительной сложностью — необходимо приближенно брать множество интегралов. Фактор граф значительно ускоряет вычисления, накладывая довольно мягкие в контексте этой задачи ограничения. Как и Elo TrueSkill имеет настраиваемый параметр влияния удачи. Здесь он интерпретируется как длина «цепочки умений». Эта «цепочка» состоит из упорядоченных по возрастанию фиксированных уровней игры. Если звено связывает уровни A и B, то игрок уровня A имеет 20% шанс выиграть у игрока уровня B. Чем цепь длиннее, тем больше разница между худшим и лучшим игроком и, следовательно, игра больше зависит от умения и меньше от везения.

Из других достоинств TrueSkill можно отметить поддержку «неполного матча», когда человек выходит из игры до ее завершения, поддержку команд различной численности и явное моделирование ничьих путем введения дополнительного параметра, отвечающего за ее вероятность.

Для дополнительного улучшения качества прогнозирования предлагается использовать композицию различных моделей TrueSkill. Уникальность моделей обеспечивается случайной подвыборкой данных для настройки их коэффициентов. Усреднение по оценкам множества таких моделей дает оценку, стабильно превосходящую по точности каждую из своих компонент. Улучшение наблюдалось на всех протестированных наборах данных.

В настоящее время вероятностный подход в моделировании рейтингов превосходит другие способы по точности и интерпретируемости, но требует больших вычислительных ресурсов. Исследованная модель TrueSkill решает эту проблему, используя фактор-граф для кэширования промежуточных результатов вычислений. Применение агрегирования моделей, настроенных на различные случайные подвыборки исходных данных дало устойчиво лучший результат.

Список литературы:

1. Википедия. Эло, Арпад – Википедия. – accessed 24-November-2015]. https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%BE_%D0%90%D1%80%D0%BF%D0%B0%D0%B4 .
2. Herbrich, Ralf. TrueSkillTM: A Bayesian Skill Rating System / Ralf Herbrich, Tom Minka, Thore Graepel // Advances in Neural Information Processing Systems 19 Ed. by B. Schölkopf, J.C. Platt, T. Hoffman. – MIT Press, 2007. – Pp. 569–576. <http://papers.nips.cc/paper/3079-trueskilltm-a-bayesian-skill-rating-system.pdf>.
3. Kschischang, F. R. Factor Graphs and the Sum-product Algorithm F. R. Kschischang, B. J. Frey, H. A. Loeliger // IEEE Trans. Inf. Theor. – 2006. – sep. Vol. 47, no. 2. – Pp. 498–519. <http://dx.doi.org/10.1109/18.910572> .
4. Minka, Thomas P. A Family of Algorithms for Approximate Bayesian Inference: Ph.D. thesis. – Cambridge, MA, USA : Massachusetts Institute of Technology, 2001. – AAI0803033.
5. Wikipedia. Bradley–Terry model – Wikipedia, The Free Encyclopedia. – 2015 – [Online; accessed 17-November-2015]. https://en.wikipedia.org/wiki/Bradley%E2%80%93Terry_model.