Bayesian Score-Based Skill Learning

Shengbo Guo · Scott Sanner · Thore Graepel · Wray Buntine

Abstract We extend the Bayesian skill rating system of TrueSkill to accommodate score-based match outcomes. TrueSkill has proven to be a very effective algorithm for matchmaking — the process of pairing competitors based on similar skill-level — in competitive online gaming. However, for the case of two teams/players, TrueSkill only learns from win, lose, or draw outcomes and cannot use additional match outcome information such as scores. To address this deficiency, we propose novel Bayesian graphical models as extensions of TrueSkill that (1) model player's offence and defence skills separately and (2) model how these offence and defence skills interact to generate score-based match outcomes. We derive efficient (approximate) Bayesian inference methods for inferring latent skills in these new models and evaluate them on three real data sets including Halo 2 XBox Live matches. Empirical evaluations demonstrate that the new score-based models (a) provide more accurate win/loss probability estimates than TrueSkill when training data is limited, (b) provide competitive and often better win/loss classification performance than TrueSkill, and (c) provide reasonable score outcome predictions with an appropriate choice of likelihood — prediction for which TrueSkill was not designed, but which can be useful in many applications.

Shengbo Guo
Xerox Research Centre Europe
E-mail: shengbo.guo@xrce.xerox.com

Scott Sanner
NICTA and ANU
E-mail: scott.sanner@nicta.com.au

Thore Graepel
Microsoft Research Cambridge
E-mail: thore.graepel@microsoft.com

Wray Buntine
NICTA and ANU
E-mail: wray.buntine@nicta.com.au

1 Introduction

In online gaming, it is important to pair players or teams of players so as to optimise their gaming experience. Game players often expect competitors with comparable skills for the most enjoyable experience; match experience can be compromised if one side consistently outperforms the other. *Matchmaking* attempts to pair players such that match results are close to being even or a draw. Hence, a prerequisite for good matchmaking is the ability to predict future match results correctly from historical match outcomes — a task that is often cast in terms of latent skill learning.

TrueSkill [5] is a state-of-the-art Bayesian skill learning system: it has been deployed in the Microsoft Xbox 360 online gaming system for both matchmaking and player ranking. For the case of two teams/players, TrueSkill, like Elo [4], is restricted to learn skills from match outcomes in terms of win, lose, or draw (WLD). While we conjecture that TrueSkill discards potentially valuable skill information carried by score-based outcomes, there are at least two arguments in favour of TrueSkill's WLD-based skill learning approach:

- WLD-based systems can be applied to any game whose outcome space is WLD, no matter what the underlying scoring system is.
- In many games, the objective is not to win by the highest score differential, but rather simply to win. In this case, it can be said that TrueSkill's skill modeling and learning from WLD outcomes aligns well with the players' underlying objective.

On the other hand, we note that discarding score results ignores two important sources of information:

- High (or low) score differentials can provide insight into relative team strengths.
- Two dimensional score outcomes (i.e., a score for each side) provide a direct basis for inferring separate offense and defense strengths for each team, hence permitting finer-grained modeling of performance against future opponents.

In this work, we augment the TrueSkill model of WLD skill learning to learn from score-based outcomes. We explore single skill models as well as separate offense/defense skill models made possible via score-based modeling. We also investigate both Gaussian and Poisson score likelihood models, deriving a novel variational update for approximate Bayesian inference in the latter case. We evaluate these novel Bayesian score-based skill-learning models in comparison to TrueSkill (for WLD outcomes) on three datasets: 14 years of match outcomes for the UK Premier League, 11 years of match outcomes for the Australian Football (Rugby) League (AFL), and three days covering 6,000+ online match outcomes in the Halo 2 XBox video game. Empirical

evaluations demonstrate that the new score-based models (a) provide more accurate win/loss probability estimates than TrueSkill (in terms of information gain) with limited amounts of training data, (b) provide competitive and often better win/loss classification performance than TrueSkill (in terms of area under the curve), and (c) provide reasonably accurate score predictions with an appropriate likelihood — prediction for which TrueSkill was not designed but important in cases such as tournaments that rank (or break ties) by points, professional sports betting and bookmaking, and game-play strategy decisions that are dependent on final score projections.

2 Skill Learning using TrueSkill

Since our score-based Bayesian skill learning contributions build on TrueSkill [5], we begin with a review of the TrueSkill Bayesian skill-learning graphical model for two single-player teams. We note that TrueSkill itself allows for matches involving more than two teams and learning team members' individual performances, but these extensions are not needed for the application domains considered in the paper.

Suppose there are n teams available for pairwise matches in a game. Let $M = \{i, j\}$ specify the two teams participating in a match and define the outcome $o \in \{team\text{-}i\text{-}win, team\text{-}j\text{-}win, draw\}$. TrueSkill models the probability $p(o|\mathbf{l}, M)$ of o given the skill level vector $\mathbf{l} \in \mathbb{R}^n$ of the teams in M, and estimates posterior distributions of skill levels according to Bayes' rule

$$p(\mathbf{l}|o, M) \propto p(o|\mathbf{l}, M)p(\mathbf{l}), \tag{1}$$

where a factorising Gaussian prior is assumed:

$$p(\mathbf{l}) := \prod_{i=1}^{n} \mathcal{N}(l_i; \mu_i, \sigma_i^2). \tag{2}$$

To model the likelihood $p(o|\mathbf{l}, M)$, each team i is assumed to exhibit a stochastic performance variable $p_i \sim \mathcal{N}(p_i; l_i, \beta^2)$ in the game [1]. From this we can model the performance differential d as an indicator function $p(d|\mathbf{p}, M) = \delta(d = p_i - p_j)$ and finally the probability of each outcome o given this differential d:

$$p(o|d) = \begin{cases} o = team\text{-}i\text{-}win : & \mathbb{I}[d > \epsilon] \\ o = team\text{-}j\text{-}win : & \mathbb{I}[d < -\epsilon] \\ o = draw : & \mathbb{I}[|d| \leq \epsilon], \end{cases} \tag{3}$$

---

[1] Note that we sometimes abuse notations on the use of p, p_i and $\mathbf{p}$. p is a probability measure; p_i and $\mathbf{p}$ represent performance variables. The meaning of them is clear from the context.
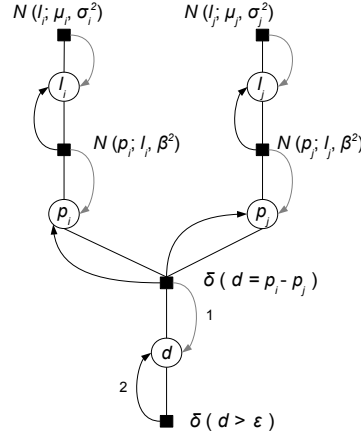
Fig. 1 TrueSkill factor graph for a match between two single-player teams with team i winning. There are three types of variables: l_i for the skills of all players, p_i for the performances of all players and d the performance difference. The first row of factors encode the (product) prior; the product of the remaining factors characterizes the likelihood for the game outcome team i winning team j. The arrows show the optimal message passing schedule: (1) messages pass along *gray* arrows from top to bottom, (2) the marginal over d is updated via message 1 followed by message 2 (which requires moment matching), (3) messages pass from bottom to top along *black* arrows.

where $\mathbb{I}[\cdot]$ is an indicator function. Then the likelihood $p(o|\mathbf{l}, M)$ in (1) can be written as

$$p(o|\mathbf{l}, M) = \int \cdots \int_{\mathbb{R}^n} \int_{-\infty}^{+\infty} p(o|d)p(d|\mathbf{p}, M) \prod_{i=1}^{n} p(p_i|l_i) \, \mathrm{d}\mathbf{p} \, \mathrm{d}d.$$

The entire TrueSkill model relevant to M is shown in the factor graph of Figure 1 with $P(o|d)$ given for the case of $o = team\text{-}i\text{-}win$. TrueSkill uses message passing to infer the posterior distribution in (1) — note that the posterior over l_i and l_j will be updated according to the match outcome while the posterior over l_k ($k \notin \{i, j\}$) will remain unchanged from the prior. An optimal message passing schedule in the TrueSkill factor graph (Figure 1) is provided in the caption; the message along arrow 2 is a step function that leads to intractability for exact inference and thus TrueSkill uses message approximation via moment matching.

TrueSkill is an efficient and principled Bayesian skill learning system. However, due to its design goals, it discards score information and does not take into account associated domain knowledge such as offence/defence skill components. Next, we propose extensions of the TrueSkill factor graph and (approximate) inference algorithms for score-based Bayesian skill learning, which address these limitations.

3 Score-based Bayesian Skill Models

In this section, we introduce three graphical models as extensions for the TrueSkill factor graph (Figure 1) to incorporate score-based outcomes in skill learning. Our first two graphical models are motivated by modeling score-based outcomes as generated by separate offence and defence skills for each team. The first generative score model uses a Poisson, which is natural model when scores are viewed as counts of scoring events. The second generative model uses a simpler Gaussian model. Our third model is a simplified version of the Gaussian model, which like TrueSkill, only models a single skill per team (not separate offence/defence skills) and places a Gaussian likelihood on the score difference, which may be positive or negative. Next we formulate each model in detail.

3.1 Offence and Defence Skill Models

In a match between two teams i and j producing respective scores $s_i \in \mathbb{Z}$ and $s_j \in \mathbb{Z}$ for each team, it is natural to think of s_i as resulting from i's offence skill $o_i \in \mathbb{R}$ and j's defence skill $d_j \in \mathbb{R}$ (as expressed in any given match) and likewise for j's score as a result of j's offence skill $o_j \in \mathbb{R}$ and i's defence skill $d_i \in \mathbb{R}$. This is contrasted with the univariate skill estimates of team i's skill l_i and team j's skill l_j used in TrueSkill, which lump together offence and defence skills for each team.

Given scores s_i and s_j for teams i and j, we model the generation of scores from skills using a conditional probability $p(s_i, s_j | o_i, o_j, d_i, d_j)$. We assume that team i's score s_i depends only on o_i and d_j and likewise that team j's score s_j depends only on o_j and d_i:

$$p(s_i, s_j | o_i, o_j, d_i, d_j) = p(s_i | o_i, d_j)p(s_j | o_j, d_i). \tag{4}$$

Like TrueSkill, we assume that the joint marginal over skill priors independently factorises:

$$p(o_i, o_j, d_i, d_j) = p(o_i)p(d_j)p(o_j)p(d_i). \tag{5}$$

Given an observation of scores s_i for team i and s_j for team j, the problem is to update the posterior distributions over participating teams' offence and defence skills. According to Bayes rule and the previous assumptions, the posterior distribution over (o_i, o_j, d_i, d_j) is given by

$$\begin{aligned} p(o_i, d_i, o_j, d_j | s_i, s_j) &\propto p(s_i, s_j | o_i, d_i, o_j, d_j)p(o_i, d_i, o_j, d_j) \\ &\propto [p(s_i | o_i, d_j)p(o_i)p(d_j)] \, [p(s_j | o_j, d_i)p(o_j)p(d_i)]. \end{aligned} \tag{6}$$

Here we observe that estimating $p(o_i, d_i, o_j, d_j | s_i, s_j)$ factorises into the two independent inference problems:

$$p(o_i, d_j | s_i) \propto p(s_i | o_i, d_j)p(o_i)p(d_j), \text{and} \tag{7}$$

$$p(o_j, d_i | s_j) \propto p(s_j | o_j, d_i)p(o_j)p(d_i). \tag{8}$$

All models considered in this paper (including TrueSkill) assume Gaussian priors on team i's offence and defence skills, i.e., $p(o_i) := \mathcal{N}(o_i; \mu_{oi}, \sigma_{oi}^2)$ and $p(d_i) := \mathcal{N}(d_i; \mu_{di}, \sigma_{di}^2)$. Our objective then is to estimate the means and variances for the posterior distributions of $p(o_i, d_j | s_i)$ and $p(o_j, d_i | s_j)$. So far, the only missing pieces in this skill posterior update are the likelihoods $p(s_i | o_i, d_j)$ and $p(s_j | o_j, d_i)$ that specify how team i and j's offence and defence skills probabilistically generate observed scores. For this we discuss two possible models in the following subsections.

3.1.1 Poisson Offence/Defence Skill Model

Following TrueSkill, we model the generation of match outcomes (in our case, team scores) based on stochastic offence and defence *performances* that account for day-to-day performance fluctuations. Formally, we assume that team i exhibits offence performance $p_{oi} := \mathcal{N}(p_{oi}; o_i, \beta_o^2)$ and defence performance $p_{di} := \mathcal{N}(p_{di}; d_i, \beta_d^2)$. With these performances, we model team i's score s_i as generated from the following process: team i's offence performance p_{oi} promotes the scoring rate while the defence performance p_{dj} inhibits this scoring rate, the difference $p_{oi} - p_{dj}$ being the effective scoring rate of the offence against the defence.

Finally, we model the score by $s_i \sim \text{Poisson}(\lambda)$, where a requirement of a positive rate λ for the Poisson distribution requires the use of $\lambda = \exp(p_{oi} - p_{dj})$ since $p_{oi} - p_{dj}$ may be negative.[2] Likewise, one can model s_j by applying the same strategy when given $\lambda = \exp(p_{oj} - p_{di})$. We represent the resulting *Poisson-OD* model in Figure 2(P) where the joint posterior is

$$p(o_i, d_j, p_{oi}, p_{dj} | s_i) \propto p(s_i | p_{oi}, p_{dj}) p(p_{oi} | o_i) p(p_{dj} | d_j) p(o_i) p(d_j),$$
$$p(o_j, d_i, p_{oj}, p_{di} | s_j) \propto p(s_j | p_{oj}, p_{di}) p(p_{oj} | o_j) p(p_{di} | d_i) p(o_j) p(d_i).$$

We are only interested in the posterior distributions of o_i, d_j and o_j, d_i given s_i and s_j, respectively. Thus, we integrate out the latent performance variables to obtain the desired posteriors

$$p(o_i, d_j | s_i) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(o_i, d_j, p_{oi}, p_{dj} | s_i) \mathrm{d}p_{oi} \mathrm{d}p_{dj},$$
$$p(o_j, d_i | s_j) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(o_j, d_i, p_{oj}, p_{di} | s_j) \mathrm{d}p_{oj} \mathrm{d}p_{di}.$$

Like TrueSkill, we use Bayesian updating to update beliefs in the skill levels of both teams in a pairwise match based on the score outcome, thus leading to an online learning scheme. Posterior distributions are approximated to be Gaussian and used as the priors in order to learn each team's skill for the next match. Approximate belief updates via variational Bayesian inference in this model will be covered in Section 4.2.

---

[2] This exponentiation of $p_{oi} - p_{dj}$ may seem to be made only to ensure model correctness, but we show experimentally that it has the benefit of allowing the Poisson model to accurately predict scores in high-scoring games even when team skills are very close (and hence $p_{oi} - p_{dj} \approx 0$).
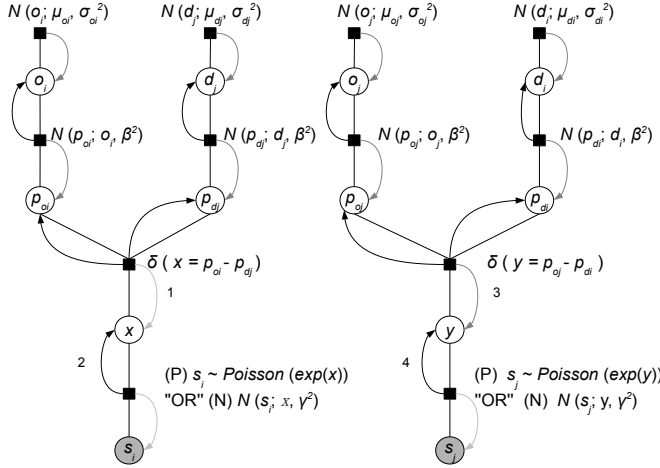
Fig. 2 The Poisson-OD (P) and Gaussian-OD (N) variants of TrueSkill factor graph for skill update of two teams based on the match score outcome (Left: modeling s_i; Right: modeling s_j). Note that the Poisson-OD and Gaussian-OD graphical models are merged due to limited space. Note also that the score observation factors use the Poisson distribution for the Poisson-OD model and the normal distribution for the Gaussian-OD model. The shaded variables are the observed ones. For each team i, it is characterized by offence skill o_i (the offence skill of team i) and defence skill d_i (the defence skill of team i). Given s_j for team j, the posterior distributions over (o_i, d_j) are inferred via message passing.

3.1.2 Gaussian Offence/Defence Skill Model

An alternative to the previous Poisson model is to model $s_i \in \mathbb{R}$ and assume it is generated as $s_i \sim \mathcal{N}(\mu, \gamma^2)$, where $\mu = p_{oi} - p_{dj}$. One can similarly model s_j by applying the same strategy when given $\mu = p_{oj} - p_{di}$. We note that unlike the Poisson model, μ can be negative here so we need not exponentiate it. While this allows us to directly model match outcomes that allow negative team scores (c.f., Halo2 as discussed in Section 5.1), it is problematic for other match outcomes that only allow non-negative team scores. One workaround would be to introduce a truncated Gaussian model to avoid the problem of assigning non-zero probability to negative scores, but we avoid this complication in exchange for the simple and exact updates offered by a purely Gaussian model.

We show the resulting *Gaussian-OD* model in Figure 2(N), which differs from our proposed Poisson model only in modeling the observed score s_i (s_j) for team i (j) given the univariate performance difference variable x (y). In this model, all messages passed during inference are Gaussian, allowing for efficient and exact belief updates.
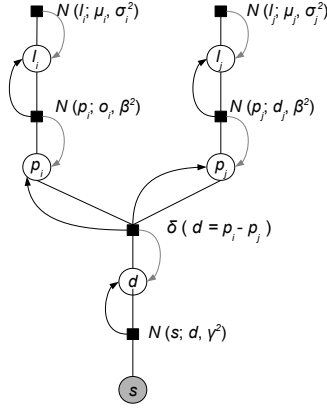
Fig. 3 Gaussian-SD model for skill learning from score differences. Both team i and team j are characterized by skill level l_i and l_j, respectively. The shaded variable s ($s = s_i - s_j$) denotes the score difference between s_i and s_j. Bayesian inference for the posterior skill level distributions has a closed-form solution.

3.2 Gaussian Score Difference (SD) Model

Again assuming $s_i \in \mathbb{R}$ and $s_j \in \mathbb{R}$, algebra for the performance means in Figure 2(N) gives:

$$s_i = p_{oi} - p_{dj}, \qquad s_j = p_{oj} - p_{di}. \tag{9}$$

This implies

$$
\begin{aligned}
s_i - s_j &= (p_{oi} - p_{dj}) - (p_{oj} - p_{di}) \\
&= \underbrace{(p_{oi} + p_{di})}_{p_{li}} - \underbrace{(p_{oj} + p_{dj})}_{p_{lj}},
\end{aligned} \tag{10}
$$

which is like modeling the score difference with performance expressions p_{li} and p_{lj} of respective univariate skill levels, l_i and l_j. Motivated by (9), we propose a score difference (SD) Gaussian model that uses a likelihood model for the observed difference $s := s_i - s_j$ specified as $s \sim \mathcal{N}(p_{li} - p_{lj}, \gamma^2)$ as shown in Figure 3.

4 Skill and Win Probability Inference

We infer skill distributions in all proposed models via online Bayesian updating. While exact inference in the purely Gaussian models can be achieved by solving linear systems, Bayesian updating provides an efficient (also exact) incremental learning alternative. Equations for Bayesian updates and win probability inference are model-dependent and presented below.

4.1 Inference in TrueSkill

Bayesian update: The Bayesian update equations in the TrueSkill model (Figure 1) are presented in [5].
Win probability: Given skill levels of team i and j, $l_i \sim \mathcal{N}(l_i; \mu_i, \sigma_i^2)$ and $l_j \sim \mathcal{N}(l_j; \mu_j, \sigma_j^2)$, we first compute the distribution over performance difference variable d, and get $d \sim \mathcal{N}(d; \mu_d, \sigma_d^2)$ with $\mu_d = \mu_i - \mu_j$ and $\sigma_d^2 = \sigma_i^2 + \sigma_j^2 + 2\beta^2$. The winning probability of team i is given by the probability $p(d > 0)$ defined as

$$p(d > 0) = 1 - \Phi\left(\frac{-\mu_d}{\sigma_d}\right), \tag{11}$$

where $\Phi(\cdot)$ is the normal CDF.

4.2 Inference in Poisson-OD Model

Bayesian update: Some of the update equations in the Poisson-OD model (Figure 2(P)) have been presented in [5], with the exception of the marginal distribution over x and the message passing from the Poisson factor to x. Given a prior Gaussian distribution over x, $\mathcal{N}(x; \mu, \sigma^2)$, we next demonstrate how to update the belief on x when observing team i's score s_i.

By the sum-product algorithm [7], the marginal distribution of x is given by a product of messages

$$p(x|s_i) = m_{\delta \to x}(x) m_{s_i \to x}(x). \tag{12}$$

To avoid cluttered notation, let us use $m_1(x)$ to represent $m_{\delta \to x}(x) = \mathcal{N}(x; \mu, \sigma^2)$, i.e., the message passing from the factor $\delta(\cdot)$ to x, and $m_2(x)$ for $m_{s_i \to x}(x) = Poisson(s_i; \exp(x))$, i.e., the message passing from the Poisson factor to x (c.f., messages labeled 1 and 2 in Figure 2(P)). Due to the multiplication of $m_1(x)$ and $m_2(x)$, the exact marginal distribution of $p(x|s_i)$ is not Gaussian, which makes exact inference intractable. To maintain a compact representation of offence and defence skills, one can approximate $p(x|s_i)$ within a variational Bayes framework by choosing a Gaussian distribution $q(x)^* : \mathcal{N}(x; \mu_{\text{new}}, \sigma_{\text{new}}^2)$ that minimizes the KL divergence between $p(x|s_i)$ and $q(x)$, i.e.,

$$q(x)^* = \underset{q(x)}{\arg\min} \, \text{KL}\left[q(x)||p(x|s_i)\right]. \tag{13}$$

We derive a fixed-point approach for optimizing $q(x)$ [12] and describe this approach below.
Minimizer $q(x)$ **for** $\textbf{KL}(q(x)||p(x|s_i))$**:** We first expand the KL-divergence into its definition:

$$\begin{aligned}
\text{KL}\left(q(x)||p(x|s_i)\right) &= \int q(x) \log\left(\frac{q(x)}{p(x|s_i)}\right) dx \\
&= -\log\sqrt{2\pi e \sigma_{new}^2} - E_{x \sim q(x)} \log\left(p(x|s_i)\right), \tag{14}
\end{aligned}$$

where $p(x|s_i)$ is the posterior probability of x when observing the score s_i. Since $q(x)$ is Gaussian and the posterior has convenient Gaussian parts, manipulation of this yields an equation for μ_{new} and σ^2_{new} that can be solved using an iterative fixed-point approach:

Lemma 1 *Values for μ_{new} and σ^2_{new} minimizing $KL\left(q(x)||p(x|s_i)\right)$ satisfy*

$$\mu_{new} = \sigma^2\left(s_i - e^\kappa\right) + \mu,$$
$$\sigma^2_{new} = \frac{\sigma^2}{1 + \sigma^2 e^\kappa}, \tag{15}$$

where

$$\kappa = \log\left(\frac{\mu + s_i\sigma^2 - 1 - \kappa + \sqrt{(\kappa - \mu - s_i\sigma^2 - 1)^2 + 2\sigma^2}}{2\sigma^2}\right). \tag{16}$$

Proof The second term in (14) is evaluated using Bayes Theorem, $p(x|s_i) = p(s_i|x)p(x)/p(s_i)$. The term in $\log p(s_i)$ can be dropped because it is constant with respect to μ_{new} and σ^2_{new}. The term $E_{x\sim q(x)}[\log p(s_i|x)]$ is found by expanding the Poisson distribution and noting $E_{x\sim p(x)}[\exp(x)] = \exp(\mu + \sigma^2/2)$ (see the Supplemental material[3] for derivation). Thus it becomes

$$s_i\mu_{\text{new}} - \exp(\mu_{\text{new}} + \sigma^2_{\text{new}}/2) - \log(s_i!) . \tag{17}$$

The term $E_{x\sim q(x)}[\log p(x)]$ according to the derivation in the Supplemental material becomes

$$-\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\sigma^2_{new} + \mu^2_{new} - 2\mu\mu_{new} + \mu^2\right) . \tag{18}$$

Plugging (17) and (18) into (14) gives

$$\underset{q(x)}{\arg\min}\, \text{KL}\left(q(x)||p(x|s_i)\right) \equiv \underset{q(x)}{\arg\min} -\log\sqrt{2\pi e\sigma^2_{new}} -$$

$$\left(\underbrace{s_i\mu_{\text{new}} - \exp(\mu_{\text{new}} + \sigma^2_{\text{new}}/2) - \log(s_i!)}_{E_{x\sim q(x)}(\log p(s_i|x))}\right.$$

$$\left.\underbrace{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\sigma^2_{new} + \mu^2_{new} - 2\mu\mu_{new} + \mu^2\right)}_{E_{x\sim q(x)}(\log p(x))}\right).$$

To find the minimizer $q(x)$, we calculate the partial derivatives of $\text{KL}\left(q(x)||p(x|s_i)\right)$ w.r.t. μ_{new} and σ_{new}, and set them to zero, leading to

$$\mu_{\text{new}} = \sigma^2\left(s_i - \exp\left(\mu_{\text{new}} + \frac{\sigma^2_{\text{new}}}{2}\right)\right) + \mu,$$
$$\sigma^2_{\text{new}} = \frac{\sigma^2}{1 + \sigma^2\exp(\mu_{\text{new}} + \frac{\sigma^2_{new}}{2})} .$$

---

[3] Available at `http://users.cecs.anu.edu.au/~sguo/sbsl_ecml2012_final_supple.pdf`

Summing the first plus half the second of these equations, and defining $\kappa = \mu_{\text{new}} + \sigma^2_{\text{new}}/2$ yields the equation for κ of

$$\kappa = \mu + \sigma^2(s_i - \exp(\kappa)) + \frac{\sigma^2}{2(1 + \sigma^2 \exp(\kappa))}, \tag{19}$$

and one gets (15) in terms of κ.

We convert (19) by solving for $\exp(\kappa)$ as it appears on the right-hand side. This yields a quadratic equation, and we take the positive solution since $\exp(\kappa)$ must be non-negative (see the Supplemental material). The result gives us (16).

We can use (16) as a fixed-point rewrite rule. For a given μ and σ^2 together with an initial value of κ, one iterates (16) until convergence. Empirically, this happens within 2-3 iterations. With convergence, we substitute the fixed-point solution into (15) to get the optimal mean and variance for $q(x)^*$.

Win probability: Suppose we are given the offence and defence skills for team i and j, we can estimate the distributions over performance difference variables of x and y (c.f., Figure 2), and compute the Poisson parameters for s_i and s_j by using $\lambda_i = \exp(x)$ and $\lambda_j = \exp(y)$. To compute the winning probability of team i, i.e., $p(s_i > s_j)$, we first construct a new variable $s = s_i - s_j$, the difference variable between two Poisson distributions, which proves to be a Skellam distribution in [10]. Thus, we can compute the win probability of $P(s > 0)$ of team i, according to the probability mass function for the Skellam distribution

$$P(s = k; \lambda_i, \lambda_j) = e^{-(\lambda_i + \lambda_j)} \left(\frac{\lambda_i}{\lambda_j}\right)^{k/2} I_{|k|}\left(2\sqrt{\lambda_i \lambda_j}\right),$$

where $I_k(z)$ is the modified Bessel function of the first kind given in [1]. We approximated $P(s > 0, \lambda_i, \lambda_j)$ with $\sum_{k=1}^{n} P(s = k; \lambda_i, \lambda_j)$ using $n = 100$ since $P(s = k; \lambda_i, \lambda_j) \approx 0$ for all of our experiments when $k > 100$.

4.3 Inference in Gaussian-OD Model

Bayesian update: In the Gaussian-OD model (Figure 2(N)), all messages are Gaussian so one can compute the belief update in closed-form as follows

$$
\begin{aligned}
\pi_{o_i} &= \frac{1}{\sigma^2_{o_i}} + \frac{1}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma^2_{d_j}}, &\qquad \tau_{o_i} &= \frac{\mu_{o_i}}{\sigma^2_{o_i}} + \frac{s_i + \mu_{d_j}}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma^2_{d_j}}, \\
\pi_{d_j} &= \frac{1}{\sigma^2_{d_j}} + \frac{1}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma^2_{o_i}}, &\qquad \tau_{d_j} &= \frac{\mu_{d_j}}{\sigma^2_{d_j}} + \frac{\mu_{o_i} - s_i}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma^2_{o_i}},
\end{aligned} \tag{20}
$$

where μ_{o_i} and σ_{o_i} are the mean and standard deviation of the prior offence skill distribution of team i, $\pi_{o_i}(\pi_{d_j}) = \frac{1}{\sigma^2_{post}}$ and $\tau_{o_i}(\tau_{d_j}) = \frac{\mu_{post}}{\sigma^2_{post}}$ are the precision

and precision-adjusted mean for the posterior offence (defence) skill distribution of team i (j). Likewise, one can derive the update equations for team j's offence skill o_j and team i's defence skill d_i.

Win probability: To compute the probability of team i winning vs team j, we first use message passing to estimate the normally distributed distributions for score variables s_i and s_j, and then compute the probability that $s_i - s_j > 0$, i.e., team i's score is larger than team j's. Given $s_i \sim \mathcal{N}(s_i; \mu_{si}, \sigma_{si}^2)$ and $s_j \sim \mathcal{N}(s_j; \mu_{sj}, \sigma_{sj}^2)$, we can compute the winning probability of team i by

$$p(s > 0) = 1 - \Phi\left(\frac{-(\mu_{si} - \mu_{sj})}{\sigma_{si}^2 + \sigma_{sj}^2}\right). \tag{21}$$

4.4 Inference in Gaussian-SD Model

Bayesian update: In the Gaussian-SD model (Figure 3), all messages are Gaussian so we can again derive the update for the single team skills l_i and l_j in closed-form as follows:

$$\pi_{l_i} = \frac{1}{\sigma_{l_i}^2} + \frac{1}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma_{l_j}^2}, \qquad \tau_{l_i} = \frac{\mu_{l_i}}{\sigma_{l_i}^2} + \frac{(s_i - s_j) + \mu_{l_j}}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma_{l_j}^2}, \tag{22}$$

$$\pi_{l_j} = \frac{1}{\sigma_{l_j}^2} + \frac{1}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma_{l_i}^2}, \qquad \tau_{l_j} = \frac{\mu_{l_j}}{\sigma_{l_j}^2} + \frac{\mu_{l_i} - (s_i - s_j)}{\beta_1^2 + \beta_2^2 + \gamma^2 + \sigma_{l_i}^2}, \tag{23}$$

where μ_{l_i} (μ_{l_j}) and σ_{l_i} (σ_{l_j}) are the mean and standard deviation of team i's (team j's) prior skill distribution, π_{l_i} (π_{l_j}) and τ_{l_i} (τ_{l_j}) are the precision and precision adjusted mean for team i's (team j's) posterior skill distribution.

Win probability: To estimate the winning probability of team i for a match with team j, one can first use message passing to estimate the normally distributed score difference variable s, and then compute the winning probability of team i by

$$p(s > 0) = 1 - \Phi\left(\frac{l_i - l_j}{\sigma_i^2 + \sigma_j^2 + 2\beta^2}\right), \tag{24}$$

where l_i and σ_i are the mean and standard deviation for team i's skill level, and β the standard deviation of the performance variable.

5 Empirical Evaluation

5.1 Data Sets

Experimental evaluations are conducted on three data sets: Halo 2 XBox Live matches, Australian Football (Rugby) League (AFL)and UK Premier League

(UK-PL)[4]. The Halo 2 data consists of a set of match outcomes comprising 6227 games for 1672 players. We note there are negative scores for this data, so we add the absolute value of the minimal score to all scores to use the data with all proposed models.

The training and testing settings are described as follows. For Halo 2 [5], the last 10% of matches are used for testing, and we use different proportions of the first 90% of data for training. There are 8 proportions used for training, ranging from 10% to 80% with an increment of 10%, and 90% is not used for training due to cross validation. To cross validate, we sample the data and run the learning 20 times at each proportion level to get standard error bars. Note that there are some players in the testing games who are not involved in any training data sets, particularly when small proportion of training data set is selected (e.g., the first 10 percent games); we remove these games in the testing set when reporting performances for all models.

For both UK-PL and AFL datasets, cross validation is performed by training and testing for each year separately (14 years for UK-PL, and 11 years for AFL). For these two datasets, we test the last 20% percent of matches in each year, with the training data increasing incrementally from 10% to 80% of the initial matches.

5.2 Evaluation Criteria

5.2.1 Information Gain

The first criterion we use to evaluate different approaches is *information gain*, which is proposed in the *Probabilistic Footy Tipping Competition*[6]: if a predictor assigns probability p to team i winning, then the score (in "bits") gained is $1 + \log_2(p)$ if team i wins, $1 + \log_2(1-p)$ if team i loses, $1 + (1/2) \log_2(p(1-p))$ if draw happens. In Section 4, we showed how to compute the win probability p for each model.

5.2.2 Win/no-Win Prediction Accuracy

While information gain provides a sense of how well the models fit the data, it is also interesting to see how accurate the models were at predicting match outcomes in terms of win/no-win (e.g., loss/draw). To compare classification performance of each model, we report the win/not winning prediction accuracy in terms of area under the curve (AUC) for the games with a win or loss outcome.

---

[4] http://www.football-data.co.uk/englandm.php

[5] Credit for the use of the Halo 2 Beta Data set is given to Microsoft Research Ltd. and Bungie.

[6] Refer to http://www.csse.monash.edu.au/~footy/

5.2.3 Score Prediction Error

We evaluate the score prediction accuracy for Poisson-OD and Gaussian-OD models for *each* team in terms of the mean absolute error (MAE). Note that we must omit the Gaussian-SD model since it can only predict score differences rather than scores.

5.3 Results

Experimental results are reported according to the parameter configurations shown in Table 1. All results are presented in Figure 4 and discussed below.

Table 1 Parameter settings. Priors on offence/defence skills: $\mathcal{N}(\mu_0, \sigma_0^2)$ with $\mu_0 = 25$ and $\sigma_0 = 25/3$. Performance variance: β, β_o, β_d.

| Model	Parameter (ϵ, γ empirically estimated)
TrueSkill	$\beta = \sigma_0/2$, ϵ: draw probability
Poisson-OD	$\beta_o = \beta_d = \sigma_0/2$
Gaussian-OD	$\beta_o = \beta_d = \sigma_0/2$, γ: score variance
Gaussian-SD	$\beta = \sigma_0/2$, γ: score difference variance

5.3.1 Information Gain

For relatively small amounts of training data (10% – 30%), the Gaussian models (OD and SD) statistically significantly outperform TrueSkill and Poisson-OD in terms of win/loss probability accuracy. On all data sets except AFL, the Gaussian models perform comparably to TrueSkill for larger amounts of training data. Gaussian-OD statistically significantly outperforms Gaussian-SD for Halo 2, indicating that separate offence/defence modeling helps.

5.3.2 Win/no-Win Prediction Accuracy

In terms of win/no-win prediction accuracy, the Gaussian-OD model generally provides the best average AUC, followed by Gaussian-SD, then TrueSkill (with the exception of cases for Halo 2 with more than 40% training data where TrueSkill performs best), then Poisson-oD. Again, we see that the separate offence/defence skill modeling of Gaussian-OD gives it a performance edge over the combined skill model of Gaussian-SD.
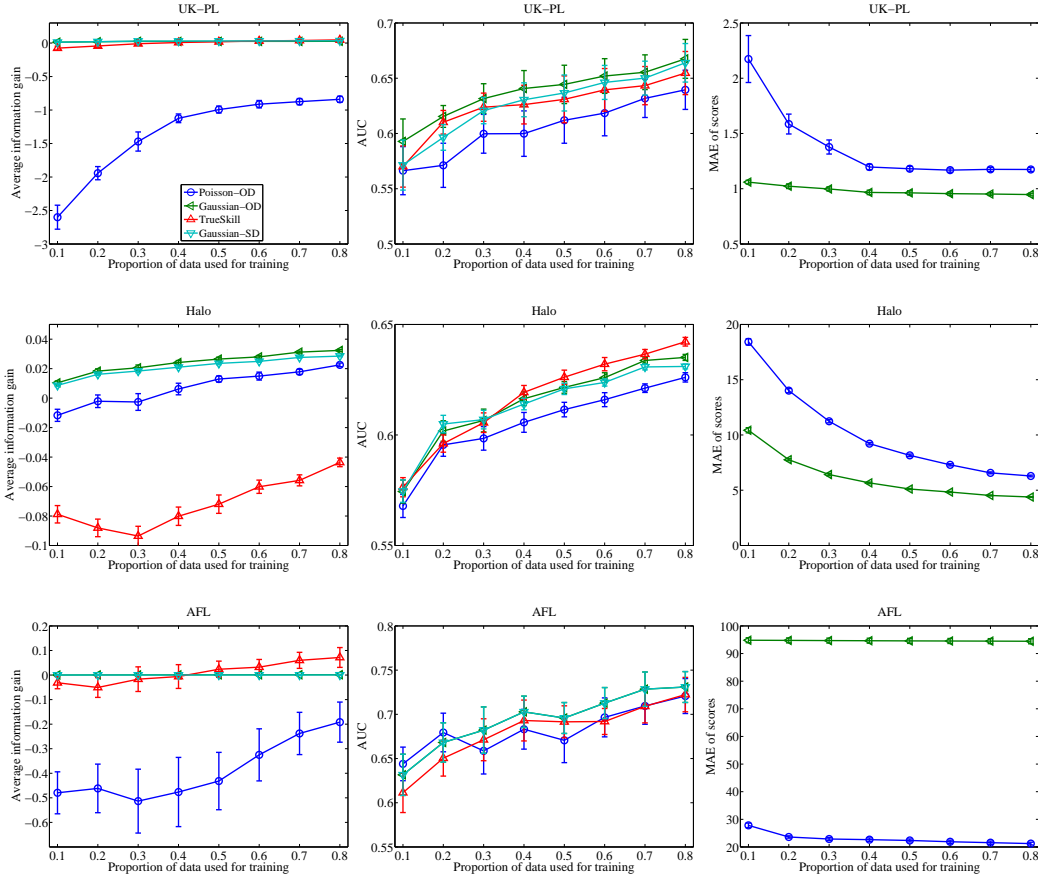
Fig. 4 Results on the UK-PL, Halo, and AFL datasets evaluated using information gain (left column), win/loss prediction accuracy in term of the area of the curve (AUC) (middle column), and score prediction error (right column). Error bars indicate 95% confidence intervals.

5.3.3 Score Prediction Errors

As shown in the third column in Figure 4, Gaussian-OD predicts more accurate scores on the UK-PL and Halo datasets, while Poisson-OD is more accurate for the AFL dataset. This can be explained by a simple skill analysis — the learned skills on the UK-PL dataset tend to show a larger variance (for all models), whereas the learned skills on the AFL dataset show little variance (for all models except Gaussian-SD). Thus, the use of an exponentiated scoring rate in the Poisson-OD model would seem to amplify these small performance differences in learned AFL skills to make more accurate score predictions on AFL data. This amplification appears to hurt the Poisson-OD model on the lower-scoring UK-PL and Halo dataset (the mean score for the AFL data is 95.4 vs 42.7 and 1.3 respectively for the Halo 2 and UK-PL data).

6 Related Work

Skill rating dates at least as far back as the Elo system [4], the idea of which is to model the probability of the possible game outcome as a function of the two players' skill levels. Players' skill levels are updated after each game in a way such that the observed game outcome becomes more likely and the summation of players' ratings remains unchanged.

The Elo system cannot handle the case when three or more teams participate in one match, a disadvantage addressed by TrueSkill [5]. Further extensions of TrueSkill incorporate time-dependent skill modeling for historical data [3].

In [2], the authors model and learn the correlation between all players' skills when updating skill beliefs, and develop a method called "EP-Correlated", contrasted with the independent assumption on players' skills (EP-Independent). Empirically, EP-Correlated outperforms EP-Independent on professional tennis match results; this suggests modeling correlations in extensions of the score-based learning presented here.

These skill learning methods all share a common feature that they are restricted to model WLD only and have to discard meaningful information carried with scores. While we proposed score-based extensions of TrueSkill in this work; it remains to incorporate other extensions motivated by this related work.

Score modeling has been studied since the 1950s [15] [16] [11] [14] [13]; one of the most popular score models is the Poisson model, first presented in [15], and this work continues to the present [13]. Other commonly used score models are based on normal distributions [11]. However, it appears that most score-based models do not distinguish offence and defence skills of each team and the results here indicate that such separate offence/defence skill models can perform better than univariate models with limited data.

7 Conclusion

We proposed novel score-based, online Bayesian skill learning extensions of TrueSkill that modeled (1) player's offence and defence skills separately and (2) how these offence and defence skills interact to generate scores. Overall these new models — and Gaussian-OD (using a separate offence/defence skill model) in particular — show an often improved ability to model winning probability and win/loss prediction accuracy over TrueSkill, especially when the amount of training data is limited. This indicates that there is indeed useful information in score-based outcomes that is ignored by TrueSkill and that separate offence/defence skill modeling does help (c.f. the performance of Gaussian-OD vs. Gaussian-SD). Furthermore, these new models allow the prediction of scores (unlike TrueSkill), with the Poisson-OD model and its variational Bayesian update derived in Section 4.2 performing best on the high-scoring AFL data. Altogether, these results suggest the potential advantages

of score-based Bayesian skill learning over state-of-the-art WLD-based skill learning approaches like TrueSkill.

Future research could combine the proposed models with related work that models home field advantage, time-dependent skills, multi-team games, and correlated skills to utilise score-based outcomes.

References

1. M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables.* Dover Publications, New York, 1974.
2. A. Birlutiu and T. Heskes. Expectation propagation for rating players in sports competitions. In *ECML-PKDD*, volume 4702 of *LNCS*, pages 374–381. Springer, 2007.
3. P. Dangauthier, R. Herbrich, T. Minka, and T. Graepel. Trueskill through time: Revisiting the history of chess. In *NIPS*, pages 337–344. MIT Press, Cambridge, MA, 2008.
4. A. E. Elo. *The rating of chess players: past and present.* Arco Publishing, New York, 1978.
5. R. Herbrich, T. Minka, and T. Graepel. Trueskill^{TM}: A Bayesian skill rating system. In *NIPS*, pages 569–576, 2006.
6. D. Karlis and I. Ntzoufras. Bayesian modelling of football outcomes: using the skellam's distribution for the goal difference. *IMA Journal of Management Mathematics*, 20(2):133–145, 2009.
7. F. R. Kschischang, B. J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47(2):498–519, February 2001.
8. T. Minka. Expectation propagation for approximate bayesian inference. In *UAI*, pages 362–369. Morgan Kaufmann, 2001.
9. M. J. Moroney. *Facts from figures.* Penguin Press Science, 3rd edition, 1956.
10. J. G. Skellam. The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society: Series A*, 109(3):296, 1946.
11. M. E. Glickman, and H. S. Stern. A state-space model for football league scores. *Journal of the American Statistical Association*, 93(441):25–35, 1998.
12. M. J. Beal, and Z. Ghahramani. The Variational Bayesian EM Algorithm for Incomplete Data: with Application to Scoring Graphical Model Structures. *Proceedings of the Seventh Valencia International Meeting*: 453–464, 2002.
13. D. Karlis and I. Ntzoufras, Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference, *IMA Journal of Management Mathematics*, 20(2):133–145, 2009.
14. D. Karlis and I. Ntzoufras, Analysis of Sports Data by Using Bivariate Poisson Models, *Journal of the Royal Statistical Society: Series D*, 52(3):381-393, 2003.
15. M. J. Moroney Facts from figures, Penguin Press Science, 3rd, 1956.
16. M. J. Dixon and S. G. Coles, Modelling Association Football Scores and Inefficiencies in the Football Betting Market, *Journal of the Royal Statistical Society: Series C*, 46(2):265–280, 1997.