

ET: Legacy Development - Feature #403

Add bayesian skill rating

18.11.2013 17:47 - Spyhawk

Status:	Fixed	% Done:	100%
Priority:	Normal		
Assignee:	Spyhawk		
Category:	Mod generic		
Target version:	2.76		
OS:		Arch:	
Description			
<p>The idea is simple: Implement a new, proved metric that can be used to compare skill of players over time.</p> <p>The existing XPs count measures player's score based on of specific action (kills, construction, revive, ..) but is far from evaluating all aspects of the 'skill' of a player. Also, XPs are useful to control the various skill levels, but don't give any indication of the real global performance - they basically only give a hint about the time the player has played on a specific server (xpsave).</p> <p>A 'skill rating' could be defined as a metric measuring "all the parameters of a player that help his team to win" and the only way to do that is by a probabilistic, Bayesian approach. It simply should look at the global output of a game instead of the various micro-parameters (accuracy, number of kill, ... you-name-it).</p> <p>ETPub implements such a metric with its "Player Rating", giving a normalized skill score. The TrueSkill system is a more modern algorithm that has been developed by MicroSoft for its Xbox matching service. It has the advantage over the ETPub PR that it starts very low and increases over time (like XPs), before stabilizing when the skill rating is accurate (like the ETPub PR).</p> <p>The algorithm is well documented (see Mooserware below) and several existing implementation can be found on the Internet, including this Lua implementation.</p> <p>Also, we could use that Skill Rating to develop several nice feature later (showing probability of winning a map, team balancing, or even match making with players of similar skills, ...)</p> <ul style="list-style-type: none">• http://en.wikipedia.org/wiki/TrueSkill• http://www.moserware.com/2010/03/computing-your-skill.html			
Update:			
More detail and analysis is available on the wiki here Bayesian skill rating			
Subtasks:			
Feature # 1139: Add promotion announcement with Skill Rating			New
Related issues:			
Related to Lua scripts for the Legacy mod - Feature # 181: Add some kind of u...	Fixed	29.01.2013	
Related to ET: Legacy Development - Feature # 887: Add new skill levels and e...	Invalid	10.11.2015	
Related to ET: Legacy Development - Bug # 1001: End of game stats not always ...	Confirmed	16.02.2017	
Related to ET: Legacy Development - Bug # 1010: Engine db_mode 1 isn't access...	New	27.02.2017	
Related to ET: Legacy Development - Bug # 1137: SR scoreboard should be defau...	Fixed	22.12.2018	
Related to ET: Legacy Development - Feature # 1155: Add Prestige to encourage...	New	27.12.2018	
Related to ET: Legacy Development - Bug # 1181: Disable SR scoreboard in SW a...	Fixed	20.01.2019	
Blocked by ET: Legacy Development - Feature # 309: ET:L database connectivity	Fixed	03.06.2013	
Blocked by ET: Legacy Development - Feature # 906: Create luasql lib and add ...	Fixed	13.12.2015	

Associated revisions

Revision a6dc54d6 - 24.12.2015 14:00 - Spyhawk

misc: added optional FEATURE_RATING, refs #403

Revision 0704e00b - 24.12.2015 14:00 - Spyhawk

misc: updated easybuild.sh, refs #403

Revision 41a8f57c - 24.12.2015 14:00 - Spyhawk

game: added initial skill rating code, refs #403

Revision 37037b62 - 24.12.2015 14:00 - Spyhawk

game: added skill rating partial play, refs #403

Revision e2764f28 - 24.12.2015 14:00 - Spyhawk

mod: added skill rating in scoreboard, refs #403

Revision f5e84f58 - 24.12.2015 14:00 - Spyhawk

mod: added skill rating in stats window, refs #403

Revision b218a576 - 24.12.2015 14:00 - Spyhawk

mod: added skill rating in debriefing window, refs #403

Revision d11b073a - 24.12.2015 14:00 - Spyhawk

mod: added skill rating in console stats dump, refs #403

Revision de4979d6 - 24.12.2015 14:00 - Spyhawk

game: don't update skill rating in stopwatch and LMS, refs #403

Revision d89a352e - 24.12.2015 14:00 - Spyhawk

mod: show rating scoreboard only if rating is enabled, refs #403

Revision 226ac1e1 - 24.12.2015 14:00 - Spyhawk

mod: added win probability, refs #403

Revision f837d448 - 24.12.2015 14:00 - Spyhawk

mod: send only complete rating to clients, refs #403

Revision ab09f983 - 24.12.2015 14:00 - Spyhawk

mod: send delta rating to clients, refs #403

Revision 8c155ab9 - 24.12.2015 14:00 - Spyhawk

game: extended rating algorithm with map parameter, refs #403

Revision 903247dd - 24.12.2015 14:51 - Spyhawk

game: log match result, refs #403

Revision f07e0620 - 24.12.2015 14:51 - Spyhawk

game: removed debug stats, refs #403

Revision b79909ed - 17.01.2016 07:03 - IR4T4

game: fix overwriting map vote values in CS_LEGACYINFO when g_skillRating is enabled refs #403

Revision e8657df8 - 17.01.2016 15:17 - IR4T4

mod: make g_skillRating latched. Don't execute rating code when disabled (WIP) refs #403

Revision 1005d01b - 30.08.2016 13:34 - Spyhawk

mod: do not execute rating code when disabled, refs #403

Revision 46f4d36d - 01.09.2016 19:09 - Spyhawk

game: disable g_skillrating by default, refs #403

Revision c9118341 - 01.09.2016 19:11 - Spyhawk

cmake: enabled FEATURE_RATING by default, refs #403

Revision bd8fc227 - 02.09.2016 20:58 - Spyhawk

game: fixed variable initialization, refs #403

Revision 4ddfaa84 - 02.09.2016 21:14 - Spyhawk

cgame: removed useless code, refs #403

Revision 40d67460 - 02.09.2016 21:40 - Spyhawk

cgame: fixed uninitialized pointer read, refs #403

Revision 86d98cd6 - 02.09.2016 21:40 - Spyhawk

game: added no division by zero check, refs #403

in case the check above doesn't catch it

Revision 3b57ec90 - 02.09.2016 22:00 - Spyhawk

misc: enable FEATURE_RATING in easybuild.sh too, refs #403

Revision 4b6ffab1 - 03.09.2016 09:30 - Spyhawk

game: cleanup, refs #403

Revision 83b26ab5 - 03.09.2016 22:51 - Spyhawk

cgame: stats info label alignment, refs #403

Revision aa8ba49f - 06.09.2016 12:54 - Spyhawk

game: added missing map win bias, refs #403

Revision 937d8a43 - 06.09.2016 12:54 - Spyhawk

cgame: display map prob on scoreboard, refs #403

Revision 4fba9212 - 12.09.2016 00:09 - Spyhawk

mod: do not display negative rating and cleanup, refs #403

Revision 506bf1d3 - 12.09.2016 11:13 - Spyhawk

cgame: fixed rating in stats window not updating when spec, refs #403

Revision b5ffb5b4 - 15.09.2016 19:20 - Spyhawk

cgame: hide skill rating in debriefing when disabled, refs #478 #403

Revision 75e31804 - 15.09.2016 20:31 - Spyhawk

cgame: fixed debriefing stats not displaying when rating is disabled, refs #403

Revision 845e2330 - 15.09.2016 21:20 - Spyhawk

mod: do not send any rating data when disabled, refs #403

Revision cba8e697 - 17.09.2016 13:44 - Spyhawk

mod: do not send delta rating in client userinfo, refs #403

Revision 0416a7d7 - 19.09.2016 14:27 - Spyhawk

mod: entirely removed rating from user configstring, refs #403

Revision 7ad92968 - 02.11.2016 14:07 - Aranud

cgame: fix logical dead code (Coverity) + fix some Clang warning

Revision a19f62ea - 03.11.2016 18:45 - Aranud

cgame: clean up + uncrustify

Revision 9d97cf94 - 23.01.2017 01:07 - Spyhawk

game: fixed total play time, refs #403

Revision 9d193942 - 23.01.2017 01:36 - Spyhawk

mod: use highest ranking as first award, refs #403 #894

Revision 52ee97dc - 23.01.2017 01:49 - Spyhawk

misc: added g_skillRating in default config file, refs #403

Revision efd9bdd1 - 23.02.2017 15:41 - Spyhawk

game: do not display meaningless total rating stats, refs #403

Revision 29a636ce - 25.02.2017 23:20 - Spyhawk

game: fixed individual rating update with partial play, refs #403

Revision 8681cbf4 - 27.02.2017 00:59 - Spyhawk

misc: enable FEATURE_DBMS if FEATURE_RATING is enabled, refs #403

Revision d69d1ae1 - 27.02.2017 00:59 - Spyhawk

game: use engine database for skill rating, refs #403

Revision 8f3d76a2 - 27.02.2017 02:16 - Spyhawk

game: added map rating bias and continuity correction, refs #403

Revision ff73c3e6 - 27.02.2017 02:16 - Spyhawk

game: set RO flag on Lua rating variables, refs #403

Revision b1590142 - 27.02.2017 02:16 - Spyhawk

game: set PRAGMA sync off with db_mode 2, refs #403

Revision e56cbaae - 27.02.2017 02:16 - Spyhawk

game: added shuffle by rating cmds, refs #403

Revision 03c7fff9 - 27.02.2017 02:16 - Spyhawk

ui: added shuffle by rating menu options, refs #403

Revision 4a6029ac - 27.02.2017 02:16 - Spyhawk

mod: added highest rated player award, refs #403 #894

Revision e26c6463 - 27.02.2017 02:35 - Spyhawk

game: set RO flag for time related Lua fields, refs #403

Revision 6a497024 - 28.02.2017 11:18 - Spyhawk

game: fixed logically dead code, refs #403

Revision 0b9bab5b - 28.02.2017 21:33 - Spyhawk

mod: display endgame stats float values with two digits, refs #403 #894

Revision f908a1c5 - 01.03.2017 11:31 - Spyhawk

game: skill rating send function cleanup, refs #403

Revision c11f4030 - 01.03.2017 22:46 - Spyhawk

game: win prob takes disconnected players into account, refs #403

Revision f73bb192 - 02.03.2017 14:51 - Spyhawk

game: fixed missing guid assignment and small optimization, refs #403

Revision 240a61e9 - 02.03.2017 14:54 - Spyhawk

game: code cleanup, refs #403

Revision 1eb2ab85 - 04.03.2017 01:07 - Spyhawk

game: do not display rating award if not positive, refs #403 #894

Revision 69b25451 - 12.03.2017 12:49 - Spyhawk

game: tied ranks to rating when enabled, refs #403

Revision 27ab8137 - 12.03.2017 12:49 - Spyhawk

game: enabled skill rating by default, refs #403

Revision 1bb1f37e - 14.03.2017 01:13 - Spyhawk

game: sanitize rank level and cleanup, refs #403

Revision 777f46b2 - 14.03.2017 01:33 - Spyhawk

game: capped display of max rating, refs #403

Revision a5ca3b40 - 14.03.2017 13:06 - Spyhawk

game: capped rating also in stats window and award, refs #403

Revision 48802e57 - 14.03.2017 15:51 - Spyhawk

game: update ranks at intermission, refs #403

Revision 40559e74 - 14.03.2017 21:52 - Spyhawk

game: use map bias for slightly better rating shuffle, refs #403

Revision 7eb991b0 - 27.08.2017 00:29 - Zack Middleton

game: fix compiling without FEATURE_RATING defined

Revision 5fd94e72 - 27.08.2017 05:48 - IR4T4

Merge pull request #279 from zturtleman/skill-rating

game: fix compiling without FEATURE_RATING defined refs #403

Revision b3349860 - 24.12.2018 01:27 - Spyhawk

cgame: enabled SR scoreboard by default, refs #403 #1137

Revision 10dcedc9 - 24.12.2018 15:57 - Spyhawk

cgame: adjusted map bias display on SR scoreboard, refs #403 #1137

Revision 19420cb6 - 26.12.2018 19:18 - Spyhawk

cgame: display scoreboard type in infotext area, refs #403 #1137

Revision e23d2622 - 26.12.2018 19:31 - Spyhawk

cgame: fixed help string overlapping chat area, refs #403 #1137

Revision a282c2d3 - 30.12.2018 17:41 - Spyhawk

cgame: removed scoreboard string to avoid overlapping, refs #403 #1137

There isn't space anywhere!

Revision 38999eb0 - 30.12.2018 18:11 - Spyhawk

cgame: display scorboard mode in help text, refs #403 #1137

Revision 711edfc7 - 31.12.2018 01:22 - Spyhawk

cgame: hide scoreboard help text in intermission, refs #403 #1137

Revision 4b6efa98 - 14.01.2019 19:04 - Spyhawk

game: ensure SR db is deinit only once in intermission, refs #403

Revision 2f242021 - 21.01.2019 13:15 - Spyhawk

cgame: fixed default scoreboard with g_SkillRating 1, refs #403 #1137

Revision 15214ace - 21.01.2019 16:47 - Spyhawk

cgame: disabled SR scoreboard in SW and LMS mode, refs #403 #1181

Revision 10068c80 - 18.04.2019 18:55 - Spyhawk

mod: only send delta rating in intermission, refs #403 #1127

History

#1 - 18.11.2013 18:21 - Spyhawk

- Description updated

#2 - 19.11.2013 13:59 - Spyhawk

- Assignee set to Spyhawk

#3 - 08.01.2015 15:24 - Spyhawk

- Description updated

#4 - 02.09.2015 21:50 - Spyhawk

- Description updated

#5 - 02.09.2015 21:52 - Spyhawk

- Description updated

Added link to some documentation in the wiki here: [Bayesian skill rating](#)

#6 - 02.09.2015 22:01 - Spyhawk

- Category deleted (Lua scripts)

#7 - 03.09.2015 16:05 - Dragonji

I'd say it would be nice to have it in the mod itself so it's easy to integrate it to current stats that are already in the game. It can't be nicely integrated in this aspect if it's a Lua script.

#8 - 03.09.2015 16:26 - Spyhawk

Dragonji wrote:

I'd say it would be nice to have it in the mod itself so it's easy to integrate it to current stats that are already in the game. It can't be nicely integrated in this aspect if it's a Lua script.

You are right, and that is actually the way I'm implementing it. The algorithm itself and integration in the mod are mostly done, but I'm still wondering if the loading/writing of data itself would be better done in Lua.

Can't change the category of the ticket though, so it's stuck to "Lua scripts" for the time being.

#9 - 04.09.2015 16:10 - Spyhawk

- Project changed from Lua scripts for the Legacy mod to ET: Legacy Development

- Category set to Mod generic

#10 - 12.09.2015 19:59 - Spyhawk

- Related to Feature #181: Add some kind of user management to the mod for online servers added

#11 - 18.11.2015 17:58 - IR4T4

- Related to Feature #887: Add new skill levels and expand military ranks like NQ Mod added

#12 - 23.11.2015 22:56 - Spyhawk

- Target version changed from ALL to 2.75

#13 - 24.11.2015 20:25 - Spyhawk

- Blocked by Feature #309: ET:L database connectivity added

#14 - 27.11.2015 20:13 - Harlekin

Would it be possible to let the server adjust some settings for a team based on team rating? Eg let them spawn a little faster, give them extra ammo on spawn, faster recharge of ammo/health cabinets, longer spawnprotection or some kind of other support for a very underpowered team.

#15 - 27.11.2015 20:32 - Spyhawk

Harlekin wrote:

Would it be possible to let the server adjust some settings for a team based on team rating? Eg let them spawn a little faster, give them extra ammo on spawn, faster recharge of ammo/health cabinets, longer spawnprotection or some kind of other support for a very underpowered team.

Anything is possible, but as a first step we'll include only informative rating. Note that rating takes strength of teams into account, so it's actually "fair" to lose against a stronger team. But if you beat a stronger team, your rating would be increased much more than usual.

#16 - 23.12.2015 17:11 - Spyhawk

The base TrueSkill algorithm has been implemented (using session to keep values for now), and I extended it for taking map and time parameters. However, to use this extension I'll need to collect data about played maps.

Data collected would be:

- number of time the map was played,
- number of time each side won the map,
- ~~average time before a team won the map~~ (not really necessary)

Here are a few possibilities:

- 1/ use a database through LuaSQL or the new engine DB feature;
- 2/ extend the existing mapvoteinfo.txt file system, and collect data in other gametypes (maybe apart from stopwatch);
- 3/ use a database system and move the existing mapvoteinfo data collect to a common database.

Option 3/ might be the cleanest, but 2/ will also work.

Opinion/suggestion?

#17 - 24.12.2015 16:57 - Spyhawk

- % Done changed from 0 to 70

#18 - 26.12.2015 17:38 - Spyhawk

- Blocked by Feature #906: Create luasql lib and add it to installs added

#19 - 30.08.2016 13:41 - Spyhawk

- Add missing Lua code (maybe as a WolfAdmin plugin) (code is moved to C)
- enable FEATURE_RATING (done)
- disable g_skillrating by default (done)

#20 - 02.09.2016 16:53 - Spyhawk

- Target version changed from 2.75 to 2.76

I'll rework the existing mapvoteinfo data to use a common database system with the map parameter extension.

Edit: After some thought, the code has been modified so map stats management can be done entirely on the Lua side (using CS_LEGACYINFO). This should stay separate from the mapvote data.

Only the accompanying Lua script is missing now.

#21 - 30.09.2016 20:43 - Spyhawk

The Lua script has been partially written. It works and takes connected players at the end of the match to compute the rating update. Also, players that disconnected and reconnected later in the same match are correctly handled. However, to take the performance of all disconnected players into account, we need to loop over all players that played in the match, and not only level.numConnectedClients.

This requires the use of an intermediary DB. I am unsure about the best way to implement this, as the C code needs to be adjusted but the DB stuff is for now entirely handled on the Lua side.

#22 - 29.01.2017 14:03 - Spyhawk

I'm hitting a wall with that DB-Lua stuff. I now believe the best way to do it is to implement directly in C, taking advantage of the engine DB. It also makes sense since the rating system is deeply integrated in the mod code already.

Special commands can still be done in Lua, in the form of a WolfAdmin plugin or independent script.

#23 - 16.02.2017 11:25 - Spyhawk

- Related to Bug #1001: End of game stats not always displayed in console added

#24 - 23.02.2017 11:09 - Spyhawk

- Put Highest ranking officer award on top
- Add variable in config file
- Tie ranks to rating instead of XPs (see G_CalcRank() and G_UpgradeSkill(), might require explicit rank update, adjust the big PM message somehow)
- track number of matches played in DB (not useful for rating)
- use ribbon (like etnam) (not useful for rating)
- remove meaningless tot_rating in G_printMatchInfo()
- enable g_skillrating by default
- add new award for best rated player
- add shuffleteamrating, shuffleteamrating_norestart votes
- add info commands such as /rating, /allrating (should be done in Lua)

WIP: move from Lua to C code: done!

- DB Init in g_main
 - DB DelInit in g_main.c
 - DB Deinit in g_vote.c (map, campaign, nextmap, maprestart, mapreset)
 - check users, match, map tables
 - empty match table at init
 - disable rating if failure
 - ClientConnect → retrieve rating from match or users tables, or create new one
 - ClientDisconnect → write in match table
 - End of game G_LogExit(): adjust G_UpdateSkillRating for use of match table
 - Add created and last updated time tracking
 - Fix edge case player connects/disconnect in intermission
 - Fix edge case player connects/disconnect in warmup
 - Fix bug where rating from rating_users table in warmup isn't taken over playing session (use empty rating_match instead)
 - Fix disconnected players aren't taken into account in scoreboard real time win probability
 - Retrieve Map prob, Save map prob
 - Display updated map rating in intermission scoreboard with g_skillRating 2
 - Fix now broken delta rating
 - Check log print G_Printf vs G_LogPrintf
 - Proper headers doc
 - Optimize speed with PRAGMA or single INSERT transaction for db_mode 2
 - Set RO flag on Lua rating variables
-
- db_mode 1 with memory mode (see #1010)

#25 - 27.02.2017 19:02 - Spyhawk

- Related to Bug #1010: Engine db_mode 1 isn't accessible from mod and LuaSQL added

#26 - 03.03.2017 15:06 - Spyhawk

- Status changed from New to In Progress

- % Done changed from 70 to 90

#27 - 11.03.2017 02:27 - Spyhawk

Some results from the DB data after 1 week of data collection (initial mu: 25, initial sigma: 8.33, rating = $\mu - 3 \cdot \sigma = 0$):

- 73 entries, from which 18 are bots (55 humans players)
- best/worst rated player
 - mu = 42.3, sigma = 6.17 (rating: 23.75).
 - mu = 20.6, sigma = 7.89 (rating: -3.10)
- human players with smallest sigma:
 - mu = 19.3, sigma = 6.15 (rating: 0.84)
 - mu = 42.3, sigma = 6.17 (rating: 23.75)

This shows that ranks based on rating and distributed in a purely linear manner (0-5: Private, ..., 45-50: General) are indeed viable. With the very limited number of players and initial bots ratings, it would require about 3 weeks for the rating to stabilize (lower variance). The more players with

higher rating (low uncertainty), the more the system can rate players quickly.

However, using a more log-like distribution of ranks (lower ranks requiring a lower rating, f.e. 0-2: Private, 2-5: Private Fist Class, ..., 40-50: General) might increase diversity of ranks on a server and also motivate new players to stay on the same server.

#28 - 12.03.2017 14:42 - Spyhawk

- Status changed from *In Progress* to *Feedback*

- % Done changed from 90 to 100

Ranks are now tied to rating in a linear manner when rating is enabled. Otherwise, behavior is still the same with ranks tied to skill levels and XPs.

We lose the Ranks popup as the ratings and thus ranks aren't updated in-game but right before intermission, but I don't think it is worth adding complexity just to show these popup at the start of next map.

Note: An equivalent implementation to rank popup is tracked in #1139.

#29 - 19.03.2017 21:05 - Spyhawk

- Status changed from *Feedback* to *Fixed*

#30 - 27.08.2018 22:34 - Spyhawk

For the record: An [updated version of the TrueSkill algorithm](#) has been available for a few months:

Briefly, the classic TrueSkill model makes the following assumptions:

1. Each player has a latent skill value that represents their expected contribution to a team. A player's performance in a game is a noisy sample of their skill.
2. The performance of a team is the weighted sum of the performances of its players, where the weight is the fraction of time the player spent on the team.
3. If a team's performance is greater than the other team by a certain margin, the team wins. Otherwise, the game is a draw. When learning skills from data, only the team win/loss information is used (not scores).
4. Player skills evolve over time according to a random walk. An increase or decrease in skill is assumed equally likely.
5. A player's skill in a game mode is assumed independent of their skill in all other modes.

The TrueSkill2 model modifies the classic model in the following ways:

- 1. A player's latent skill is inferred from their individual statistics such as kill and death counts, in addition to team win/loss.**
2. When a player quits or drops out in the middle of a game, it is treated as a surrender and their skill is updated as if they lost a game (regardless of actual outcome).
3. A player's skill in a game mode is assumed statistically correlated with their skill in other modes, so that when a player starts a new mode, their skill rating from other modes is borrowed.
4. The random walk of player skill is assumed biased toward increasing skill, with larger bias during the first matches a player plays in a game mode.
5. When a player is part of a squad, their performance is assumed to be better than normal.

There is no existing implementation as of yet (only the paper). Might worth a look sometimes in the future.

#31 - 22.12.2018 22:55 - Spyhawk

- Related to Bug #1137: SR scoreboard should be default when g_skillRating is enabled added

#32 - 27.12.2018 13:39 - Spyhawk

- Related to Feature #1155: Add Prestige to encourage XP reset added

#33 - 20.01.2019 21:37 - Spyhawk

- Related to Bug #1181: Disable SR scoreboard in SW and LMS added