

基于机器学习的 RTS 游戏实时胜率预测

温叶廷, 黄海于

(西南交通大学 信息科学与技术学院, 四川 成都 611756)

摘要: 游戏胜负预测可用于自适应游戏 AI 的设计, 策略层面强化学习的反馈参数等。使用 SC2LE 公开的数据集, 首先通过游戏时间、MMR 和 AMP 指标进行数据预处理, 得到质量较高的数据集; 然后使用 pysc2 解析, 提取游戏数据; 最后进行特征分析, 得到基础特征和统计特征, 完成游戏特征数据集的构建。最终采用机器学习方法 XGB 分类模型, 利用 10 次十折交叉验证法进行模型评估与优选。结果表明使用基础特性与统计特性的组合, 可以使得实时胜率预测准确率在不同匹配对局情况下均超过 80%。

关键词: AI; 游戏; 实时; 机器学习; XGB

中图分类号: TP391

文献标识码: A

DOI: 10.19358/j.issn.2096-5133.2020.03.007

引用格式: 温叶廷, 黄海于. 基于机器学习的 RTS 游戏实时胜率预测[J]. 信息技术与网络安全, 2020, 39(3): 35-39.

RTS game real-time winning rate prediction based on machine learning

Wen Yeting, Huang Haiyu

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: Game outcome prediction can be used in the design of adaptive game AI and feedback parameters for reinforcement learning at the strategic level. In this paper, the data set disclosed by SC2LE (StarCraft II Learning Environment) is used to preprocess the data through game time, MMR (matchmaking rating), and AMP (actions per minute) indicators to obtain a high-quality data set; then the pysc2 is used to analyze and extract the game data, and finally the feature analysis is carried out to obtain basic features and statistics features and complete the construction of game feature datasets. Finally, the machine learning method XGB classification model is used, and 10 times 10-fold cross-validation method is adopted for model evaluation and optimization. The results show that using the combination of basic characteristics and statistical characteristics, the real-time win rate prediction accuracy rate can exceed 80% under different matching games.

Key words: AI; game; real time; machine learning; XGB

0 引言

实时战略游戏 (RTS) 在过去十年中一直是人工智能研究领域的一个热点^[1]。由于 RTS 游戏复杂性、部分可观察和动态实时变化战局等游戏特点使得研究变得比较困难, 因此对于获胜者的预测成为了人工智能研究的一个高度相关的主题。

现有的研究主要集中在 MOBA 游戏 DOTA2 中, 比如, 文献[2]使用两种不同的方法对游戏胜负做出了预测, 第一种使用完整的赛后数据, 第二种在英雄选择阶段, 上述的两种方法是用来预测整局游戏的胜负, 而不是实时胜率; 文献[3]将赛前特征与比赛期间的特性相结合, 使用逻辑回归, 得到了实时游戏预测胜率的准确率为 71.49%。虽然目前

研究在游戏胜负预测方面已经有了一些成果, 但是由于游戏不同, RTS 与 MOBA 游戏有着很大的区别, 而被广泛认为是最具挑战性的 RTS 游戏的星际争霸 II^[4] 仍然没有得到解决。星际争霸 II 与其他的即时战略游戏一样分为双人或多人游戏。每一名玩家可以选择不同的种族 (人族 T、虫族 Z 和神族 P) 并且可以组建一支军队来击败敌人, 而且每一位玩家可能在一场游戏中同时操作很多单位。想要建立军队, 玩家就需要开采资源、建造建筑、研究科技和训练单位。最后通过摧毁敌方建筑获得胜利。

本文基于 SC2LE 公开的星际争霸 II 游戏数据集, 该数据集是天梯赛 1V1 数据, 总计 106 万场。本文通过对海量的游戏数据信息进行处理, 得到了

一些与游戏胜负相关的特征,这些特征主要分为两类:基础特性和统计特性,通过实验表明,与单独使用基础特性相比,加入统计特性能够提高实时胜率预测的准确率。

1 构建游戏数据集

1.1 数据预处理

利用游戏的回放数据去预测游戏的胜负是一种科学有效的方法。但是游戏数据集的质量好坏是影响预测结果的重要因素,因此,如何从 106 万局游戏中筛选出高质量的回放数据是至关重要的。本文构建游戏数据集基于以下原则:(1)选择高水平玩家的对局;(2)根据不同匹配对局游戏时长的均值和标准差选取游戏时间。

选取高水平玩家的对局主要使用两个指标:MMR(星际争霸 II 大区中比赛匹配等级, Match Making Ration)和 APM(玩家的每分钟操作数, Actions Per Minute)。图 1 所示为 TvZ 对局中 MMR 与 APM 的散点图,从图中可以看出两者没有相关性,而且高的 MMR 并不一定会有高的 APM,反之亦然。因此,应该将两者结合使用。

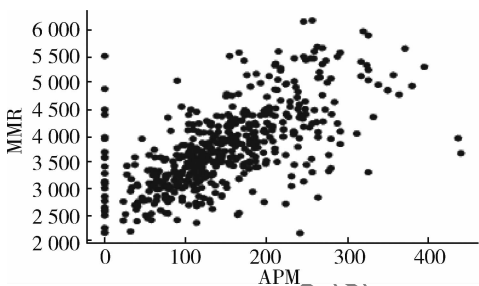


图 1 TvZ 对局的 MMR 和 APM 散点图

游戏时间的筛选也是判断游戏质量的指标,游戏时间太短或太长的对局,会造成数据偏差太大。通过计算不同匹配组合游戏时间的均值和标准差来选择游戏时长的范围。如表 1 所示,由于游戏中有 3 种不同的种族,因此会有 6 种不同的匹配对局,分别为: PvT、PvZ、TvZ、PvP、ZvZ 和 TvT。范围中的下限使用均值向下取整与标准差向上取整相减所得,上限使用均值向上取整与标准差向上取整相加所得。

表 1 不同匹配对局的均值、标准差和选择区间 (min)						
VS	PvT	PvZ	TvZ	PvP	ZvZ	TvT
均值	12.7	11.58	14.0	11.58	11.52	13.2
标准差	5.22	4.20	6.00	4.20	4.19	5.30
范围	6 ~ 19	6 ~ 17	8 ~ 20	6 ~ 17	7 ~ 15	7 ~ 20

这里将使用表 1 中的时间区间,并结合 MMR (大于 4 000)、APM(大于 150,接近专业玩家水平)进行筛选。筛选前后对局数量如表 2 所示,FTS 表示原始数据中的对局数量,ATS 表示筛选后的对局数量。

表 2 筛选前与筛选后的对局数 (万局)

	VS	PvT	PvZ	TvZ	PvP	ZvZ	TvT
FTS	26.92	23.26	8.43	18.94	11.78	16.64	
ATS	1.03	0.90	1.97	0.30	1.15	0.86	

1.2 解析数据

将筛选下来的高质量对局通过使用 pysc2 进行数据提取,在回放的每一帧都会返回 Observation,使用 ProtoBuf 进行解析,解析后的每个观察值包括以下信息:

- (1) 玩家所拥有的建筑、单位和技术;
- (2) 玩家拥有的人口情况;
- (3) 玩家所使用和拥有的资源;
- (4) 玩家所观察到的敌人单位和建筑。

1.3 特征提取

1.3.1 基本特征

从游戏回放中解析出游戏状态,在这里,首先提取出基本的特征。这些基本特征主要包括:

- (1) 资源信息:包括当前时刻的瓦斯数量和晶体矿的数量。
- (2) 人口信息:包括当前时刻人口上限、已使用的人口数和未使用的人口数。
- (3) 单位信息:包括每种单位的数量,已经训练完毕的单位数量,正在建造的数量,快要建造完毕的数量,刚开始建造的数量以及平均建造数量。
- (4) 研究科技信息:包括已经升级的科技。
- (5) 观测到的敌方玩家建筑和单位的数量。

将上面的信息进行序列化,但是由于种族不同,单位的类型数量以及研究科技的种类也不相同,因此不同的对局对应的序列长度也不一样,如表 3 所示。

表 3 不同种族对应的单位类型数量与科技研究种类

种族	T	P	Z
单位种类数量(<i>U</i>)	56	41	73
研究科技的种类(<i>R</i>)	31	28	28
建筑种类数量(<i>B</i>)	28	18	19

每场游戏的基本回放特征会被组织成一个 $N \times M1$ 大小的矩阵, N 指行数, 与游戏时间相关; $M1$ 根据使用的种族和对战种族不同而不同。表 4 所示为 $M1$ 的组织形式。不同的种族由于其单位种类和研究种类数量上的差异, 特征序列的长度也不相同, 如表 5 所示。

表 4 序列对应的每个字段的含义

特征编号	特征含义	备注
[0]	游戏帧	反映游戏时间
[1,14]	累计分数	游戏中获得的部分分值
[15,16]	两种资源类型	代表两种资源的数量
[17,19]	人口情况	分别代表人口最大值、使用的人口数和未使用的人口数
[20, R)] ¹	科技研究情况	已经研究为 1, 否则为 0
[R , $S(U)$)] ²	己方单位情况	显示单位信息
[$S(U)$, $E(U)$)] ³	敌方单位情况	可观察的敌方单位数量

注: (1) R 大小为种族的研究数量(见表 3)加 20;
(2) $S(U)$ 大小为 R 加上 $6 \times$ 种族的单位种类数量(见表 3);
(3) $E(U)$ 大小为 $S(U)$ 加上敌方单位种类数量(见表 3)。

表 5 不同匹配对局不同种族的特征序列长度

VS	种族	基础特征向量($M1$)
PvP:P	P	335
TvT:T	T	443
ZvZ:Z	Z	559
PvT:T	T	428
PvT:P	P	350
PvZ:Z	Z	527
PvZ:P	P	367
ZvT:T	T	460
ZvT:Z	Z	542

1.3.2 统计特征

除了上述的基本特征以外, 这里提出了有关动作和地图的统计信息。统计信息主要分为两类, 一类是对动作做统计, 另一类是对地图和距离的部分信息进行统计。

星际争霸 II 是一个需要有效利用游戏时间的游戏, 一般在游戏过程中分为前期、中期和后期三部分, 想要在每个时期达到预期的效果, 则采用的命令也就不同。因此, 命令的类型以及频率能够反

映出游戏的状态, 比如在前期时会大力地发展经济, 所以关于建造的命令就相对较多。

首先统计某类类型的命令频率。可以发现最常用的 6 种命令为: MOVE, BUILD, TECH, TRAIN, SIEGE 和 BURROW。而且根据执行命令是否使用资源将命令分为两大类: 一类是在执行过程中不使用资源, 称为微观命令; 否则为宏观命令。

受文献[5]的启发, 可以将命令分为三类: 控制命令、策略命令和战术命令。

动作统计信息包括上述最常用的 6 种命令, 还包括微观命令、宏观命令、控制命令、策略命令和战术命令, 共计 11 种命令。这些信息每隔 1 min 间隔统计一次。

SC2LE 中的数据集主要是收集天梯赛的对局, 天梯赛在每个版本都会有几种不同的地图。每种地图都有各自的风格, 地图的主要特点主要有补给距离、空间距离、矿区范围、矿区高地口宽度和矿区的资源数量。根据文献[6]对于游戏地图特点的分析, 可以得出地图对于玩家的胜负也是较为关键的。

除了基本的地图统计信息以外, 建筑物的摆放位置对于游戏胜负的影响很大, 好的摆放位置, 可以用来抵挡敌人^[7], 在这里通过计算同种类型建筑物的相对距离和指挥中心周围建筑与指挥中心的相对距离两个指标, 来反映出建筑物的位置信息。式(1)计算了同种建筑的相对距离, 因为游戏地图较大, 所以使用指挥中心进行分组, 将距离某个指挥中心较近的分为一组, 指挥中心的个数就是组的个数, 用 m 表示。

$$d_{ik} = \frac{1}{n} \sum_{j=1}^{n-1} |u_{i0} - u_{ij}| \tag{1}$$

其中, u_{ij} 代表第 i 类建筑中第 j 个单位的位置, n 代表某组中某种建筑的数量, d_{ik} 代表第 k 组 i 类单位的相对距离。

$$D_i = \frac{1}{m} \sum_{k=0}^{m-1} d_{ik} \tag{2}$$

D_i 为 i 类单位的平均距离。

式(3)表示的是建筑物相对于指挥中心的距离, 同样也是按照指挥中心分组计算, 一般情况下, 建筑物都是围绕着指挥中心建造的。

$$x_k = \frac{1}{s} \sum_{j=0}^{s-1} |c_k - p_j| \tag{3}$$

其中 c_k 为第 k 个指挥中心的位置, p_j 为第 k 组中第 j

个建筑的位置, s 为第 k 组的单位个数。

式(4)表示指挥中心周围建筑与指挥中心的相对距离。对于建筑物的位置摆放需要集中, 过于分散不利于防守。

$$X = \frac{1}{m} \sum_{k=0}^{m-1} x_k$$

(4)

统计特征包括动作统计信息和地图统计信息, 表 6 所示为统计特征的组织形式。

表 6 统计特征每个字段的含义

特征编号	特征含义
[0,10]	动作统计
[11,15]	地图特点
[16]	指挥中心周围建筑 与指挥中心的相对距离
[17,S(B)) ¹	同种类型建筑物的相对距离

注: $S(B)$ 大小为 17 加上建筑物类型数量(见表 3)减 1。

统计特征不同于基础特性, 其大小与匹配对局无关, 只与种族相关。如表 7, 统计特征大小用 $M2$ 表示。

表 7 统计特征序列长度

种族	P	T	Z
统计特征向量($M2$)	35	45	36

2 实验结果与分析

每场游戏的回放特征会被组织成一个 $N \times M$ 大小的矩阵, N 与回放游戏时间有关。在本文中, 采取每隔 1 min 构建一个 M , 但是在提取数据时为了不丢失动作信息是按照每 8 帧提取一次数据^[8], 这里 M 的大小是 $M1 + M2$, 例如 $M(T)_{PvT} = M1(T)_{PvT} + M2(T) = 428 + 45 = 473$ 。经过数据处理后, 得到每种不同匹配对局下的数据总条数, 对每局游戏每隔 1 min 做一次数据采样作为一个 M (游戏刚开始不做采样), 最后每一种匹配的数据总条数如表 8 所示。

表 8 不同对局下的数据条数 (万条)

VS	PvT	PvZ	TvZ	PvP	ZvZ	TvT
条数	13.4	11.9	28.4	3.8	13.6	11.5

在本节中, 将使用机器学习算法 XGB 构建模型, 将预测胜负转化为一个二分类任务。首先, 为

每一种匹配类型设计一个单独的模型, 用于预测不同对局下的实时胜率。

XGBoost^[9] 是 Gradient Boosting 方法的一个特定实现, 使用更精确的近似来找到最佳树模型。它采用了许多漂亮的技巧, 使其非常成功, 尤其是结构化数据。最重要的是, 计算二阶梯度, 即损失函数的二阶导数(类似于牛顿方法), 其提供了关于梯度方向以及如何达到损失函数最小值的更多信息和高级正则化(L1 和 L2), 提升了模型的泛化能力^[10]。

将数据集按照 6:2:2 分为训练集、验证集和测试集, 在训练过程中对样品进行了 10 次交叉验证。为了避免偏差, 对于任意对局, 样本要么在训练集中, 要么在测试集中, 但不同时存在。如表 9 所示, B 代表基本特征情况下的准确率, B + S 代表组合特征情况下的准确率。

表 9 基本特征与组合特征下的实时准确率 (%)

VS	B	B + S
PvT:P	76.53	83.35
PvZ:P	75.67	82.13
PvZ:Z	75.15	80.26
PvT:T	74.65	83.18
TvZ:T	75.39	80.17
TvZ:Z	77.65	83.89
TvT:T	73.07	80.21
PvP:P	76.98	83.19
ZvZ:Z	76.02	82.26

从上述实验结果可以看出, 本文提出的组合特性的预测效果比单独使用基础特征的效果要好。

3 结束语

游戏实时胜率预测是游戏 AI 的重要一环, 本文通过使用 SC2LE 公开的数据集构建特征数据集, 分析了不同匹配对局下使用基础特征和统计特征时不同模型的研究结果, 结果表明, 在基础特性中加入统计特性能够提升实时胜率预测的准确率。通过使用本文中的预测模型, 可以解决在游戏 AI 中使用强化学习时关于奖励稀疏的问题, 能够加快强化学习的模型训练。

参考文献

- [1] 李博. 游戏人工智能关键技术的研究[D]. 上海: 上海交通大学, 2011.
- [2] KINKADE N, JOLLA L, LIM K. Dota 2 win prediction[J]. Univ Calif, 2015, 1: 1-13.
- [3] RAVARI Y N, BAKKES S, SPRONCK P. Starcraft winner prediction[C]. Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference, 2016.
- [4] VINYALS O, EWALDS T, BARTUNOV S, et al. Starcraft ii: a new challenge for reinforcement learning[J]. arXiv preprint arXiv:1708.04782, 2017.
- [5] ONTANON S, SYNNAEVE G, URIARTE A, et al. A survey of real-time strategy game AI research and competition in StarCraft[J]. IEEE Transactions on Computational Intelligence and AI in Games, 2013, 5(4): 293-311.
- [6] BARRIGA N A, STANESCU M, BURO M. Building placement optimization in real-time strategy games[C]. Tenth Artificial Intelligence and Interactive Digital Entertainment Conference, 2014.
- [7] CERTICKY M. Implementing a wall-in building placement

in starcraft with declarative programming[J]. arXiv preprint arXiv:1306.4460, 2013.

- [8] WU H K, ZHANG J G, HUANG K Q. MSC: a dataset for macro-management in StarCraft II[J]. arXiv preprint arXiv:1710.03131, 2017.
- [9] CHEN T Q, GUESTRIN C. Xgboost: a scalable tree boosting system[C]. Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 785-794.
- [10] ZHANG D H, QIAN L Y, MAO B J, et al. A data-driven design for fault detection of wind turbines using random forests and XGboost[J]. IEEE Access, 2018, 6: 21020-21031.

(收稿日期: 2020-01-14)

作者简介:

温叶廷(1995-), 男, 硕士研究生, 主要研究方向: 即时战略游戏中的人工智能。

黄海于(1970-), 男, 硕士, 副教授, 主要研究方向: 计算机应用、人机交互。

(上接第 34 页)

- [12] 林加润, 殷建平, 张晓峰, 等. 云计算环境下安全的极限学习机外包优化部署机制[J]. 计算机与数字工程, 2019, 47(1): 157-160, 230.

(收稿日期: 2020-01-14)

作者简介:

李雪艳(1983-), 女, 硕士, 实验师, 主要研究方向: 图像识别应用、深度学习。

廖一鹏(1982-), 通信作者, 男, 博士, 主要研究方向: 图像处理、模式识别。E-mail: fzu_lyp@163.com。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所