

基于变化概率的网络选课系统公平算法

张 戈 (中国青年政治学院 计算机教学与应用中心 北京 100089)

摘要: 网络选课系统研发工作的重点和难点就是找到一种算法既能充分体现选课活动的公平性又能很好地解决大批量并发访问系统带来的网络拥堵问题。对现有选课算法进行了分析和研究,提出并详细阐述用于实时选课系统的一种新算法——基于变化概率的公平算法。该算法使得每个学生在选课活动中能拥有趋于均等的选课概率,从而使选课活动在最大程度上实现了公平性原则并极大的缓解了网络拥堵问题。

关键词: 选课算法;变化概率;公平性;网络拥堵

Algorithms of Web Course Selection-Arrangement System Based on Varied Probability

ZHANG Ge

(China Youth University for Political Sciences, Beijing 100089, China)

Abstract: The most important and difficult thing is how to find an algorithm which can resolve the equity problem and network congestion problem caused by concurrent visit to system by multitudinous students. This paper analyzes and researches these existing algorithm, and proposes new algorithms - Algorithms of Web Course Selection-Arrangement System based on probability. These algorithms make every student have equal probability on the move of course selection. Therefore, the equity and network congestion can be resolved at large.

Keywords: algorithms of Web course selection-arrangement system; varied probability; the equity problem; network congestion

当前网络选课系统已广泛应用于国内各大专院校,系统研发工作的重点和难点就是找到一种算法既能充分体现选课活动的公平性又能很好的解决大批量并发访问系统带来的网络拥堵问题,但遗憾的是目前还没有一个非常行之有效的算法可以同时解决上面两个问题。

1 现有选课系统算法分析

目前常见的算法有先到先得筛选算法、抽签筛选算法、专业优先筛选算法、分志愿筛选算法^[1]等,这些算法各有利弊。比如先到先得筛选算法,这种算法由于其简单易实现的特性被不少院校所采用。但是该算法会造成大量学生在选课初期时间段内集中访问系统进行选课,导致网络拥堵,加大系统负担,甚至

会出现系统瘫痪的情况^[2]。抽签算法在数学上满足了公平性原则,但是却忽略了学生专业、学生兴趣等多方面因素对课程的需求,没有体现真正的公平^[3]。专业优先筛选算法只适用于专业选修课,适用面狭窄不能堪当整个选课重任。分志愿筛选算法是一个顾及到学生意愿的公平算法,但是该算法只适用于同一门课程开设多个教学班的情况,也具有一定的局限性^[4]。

对选课算法的现状进行研究和分析之后,发现不同算法都存在这样或那样的问题,各个院校均采用结合自己实际情况选择一种或多种选课算法的方式来处理选课活动^[5]。在对一些学校的课程体系分析之后,发现很多学校的课程设置均存在共性,即选修课分为公共选修和专业选修两大类。公共选修课的选课活动占整个选课活动的大部分,选课对象范围最广,参与

人数最多,公选课的执行能力直接影响整个选课活动的效率,因此本文针对公共选修课提出了一种新的选课筛选算法。

2 公共选修课筛选算法设计

2.1 实时选课方式的采用

由于公共选修课面向学校所有系所有专业学生,学生的专业差异被弱化,对于这样的课程采用实时选课的方式是可行的^[6]。实时选课就是实时跟踪学生的选课情况,实时反映系统总课表的调整情况,以达到学生选课的准确无误^[7]。一般地实时选课系统应具有如下功能:学生能直观的看到全校每门课程的所有开课信息,包括真实的上课时间等,而不是时间代码,并能实时监测每门课程的被选情况;学生可直观地看到自己随着选课操作而不断变动的个人选课信息;对于每次选课操作,能实时判断选课时间和选课人数冲突并反馈选课结果给学生;选课中出现课程调整时,能在选课时使用的总课表中做出实时反应^[8]。实时选课方式与“前台选课后台筛选”的选课方式的最大区别在于其选课的实时性。在此方式下,选课系统会对学生的选课概率做出实时调整,当选课结束或者该课程达到满员时,选中学生的人数和这些学生的学号已经基本确定,可直接告知学生这一选课结果,如果需要,教务工作人员只做部分调整^[9]。教务工作人员不需要在后台对该课程的所有学生做筛选。

2.2 变化概率公平算法设计

针对实时选课,本文提出“变化概率公平筛选算法”。基于这种算法,学生选课时的选课概率会实时变化和调整,每个学生具有接近均等的选课概率,从而使选课活动的公平性达到最大化。该算法描述如下:

步骤 1:设置初始值,包括初始选课概率(即第一个参选该门课程的学生具有的选课概率)和初始随机数:初始选课概率的设置以以往该课程的选课情况为参考,具体做法为,取出上次该课程的选课筛选表,将其初始选课概率(即第一个人的选课概率)设置为 1,然后取所有学生选课概率的平均值;初始随机数 = 随机函数 × 初始选课概率;

步骤 2:在选课过程中,每有一个学生选择该课程,则随即统计当前选择该课程的参选人数 A,该学生在所有选课记录中的序号 B,当前该课程的选中人数 Z,并计算该学生的选课概率 P,选课概率 P 的计

算如公式(1)。

$$P = \frac{A - Z}{(B - Z)^{1/2}} \quad (1)$$

说明:第一,此公式中的符号“^”代表开方;第二,由于参数 B 是选课学生在当前所有选课记录(包括所有公共选修课的选课记录)中的排序序号,其值总是大于当前该课程的选中人数 Z,所以此公式总有解。例如,当前已经有 30 人进行了公选课(所有公选课)的选课活动,总选课记录为 30,一名学生在选择了《计算机网页设计》课程之后,该学生对应的参数 B 即为 31,而当前《计算机网页设计》课程的选中人数 Z 总小于参数 B,则公式总有解。

步骤 3:计算随机数 R, $R = \text{RAND}() \times \text{系数}$,系数为前面所有选课概率中的最大值;

步骤 4:对该生的选课概率 P 和随机数 R 进行比较,如果 $P > R$,则该生为选中状态,否则该生落选;

步骤 5:判断选课时间是否结束或者选中人数是否达到预置人数上限 M,两个条件有一个满足,则结束该课程的选课活动;

步骤 6:该课程选课活动结束时,统计该课程的总参选人数 S,总选中人数 T,如果 $S \leq M$,则做出调整:将所有该课的参选学生设置为选中状态,否则对 T 和 M 进行比较,并做出如下处理:

如果 $T = M$,则保留当前筛选结果,该结果为正式选课结果;

如果 $T > M$,则从选中学生中随机抽取 $T - M$ 个,将其选课结果设置为落选;

如果 $T < M$,则从落选学生中随机抽取 $M - T$ 个,将其选课结果设置为选中。

假设系统当前共有 120 人参加所有课程的选课活动(这里的当前是指同一时间,即同一时刻),有 35 人参加《计算机网页设计》课程的选课活动,该课程的预置上限人数为 30 人,其选课筛选表如表 1 所示。

对上面筛选表做选课概率的折线图,如图 1 所示。

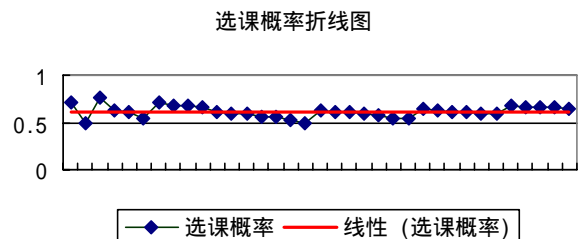


图 1 选课概率折线图

表 1 选课筛选表

选择该课程的参选人数	在总选课范围内的编号	随机数	选课概率	是否选中	选中人数
1	2	0.419054	0.707107	1	1
2	5	0.683012	0.5	0	1
3	8	0.105883	0.755929	1	2
4	12	0.30159	0.632456	1	3
5	14	0.576887	0.603023	1	4
6	18	0.558359	0.534522	0	4
7	22	0.330919	0.707107	1	5
8	25	0.617663	0.67082	1	6
9	26	0.053663	0.67082	1	7
10	28	0.152579	0.654654	1	8
11	32	0.015399	0.612372	1	9
12	35	0.461983	0.588348	1	10
13	36	0.246064	0.588348	1	11
14	39	0.556689	0.566947	1	12
15	41	0.375968	0.557086	1	13
16	45	0.44258	0.53033	1	14
17	52	0.698443	0.486664	0	14
18	55	0.586983	0.624695	1	15
19	57	0.189116	0.617213	1	16
20	59	0.59162	0.609994	1	17
21	63	0.030503	0.589768	1	18
22	67	0.225489	0.571429	1	19
23	73	0.303238	0.544331	1	20
24	76	0.606955	0.534522	0	20
25	80	0.222197	0.645497	1	21
26	83	0.237839	0.635001	1	22
27	88	0.339915	0.615457	1	23
28	92	0.507002	0.601929	1	24
29	94	0.414337	0.597614	1	25
30	95	0.637403	0.597614	0	25
31	102	0.277938	0.683763	1	26
32	107	0.485775	0.666667	1	27
33	110	0.408629	0.658586	1	28
34	111	0.388202	0.658586	1	29
35	116	0.712199	0.643268	0	29

3 公平选课算法的特点及难点分析

从选课筛选表和选课概率折线图可以看出, 本文设计的基于变化概率的公平选课算法具有以下特点。

(1) 该算法为每个选课学生计算其对应的选课概率;

(2) 该算法对选课概率做出及时合理的调整, 具体表现为每个学生的选课概率受其前面学生选中状态的影响: 当前面学生为选中状态时, 则该算法对后面学生的选课概率进行下调; 反之, 当前面学生为落选状态时, 则该算法对后面学生的选课概率进行上调;

(3) 学生的选课概率在选课概率均值 0.61 上下波动, 随着选课人数的增加, 选课概率波动幅度越来越小, 最后趋于一个数值。这一现象反映出学生选课概率的均等性, 学生选中与否与其选课先后次序无绝对关系, 从而将选课活动的公平性最大化。

(4) 在该算法下, 学生选中与否和其选课先后顺序无关, 极大的缓解了选课活动初期阶段大批量学生并发访问系统带来的网络拥堵问题, 优化了选课系统。

该算法的难点在于初始选课概率的设置, 因为该数值用以往该课程的选课情况作为基准参考, 如果上届该课程的选课情况与本次该课程的选课情况基本吻合或相近, 则初始概率比较准确, 如果上届该课程的选课情况与本次该课程的选课情况相差很远, 则初始概率偏差较大, 影响选课结果的准确性。对于这种情况, 建议的解决办法为, 在选课的步骤 5 与步骤 6 之间设置一个监视点, 由教务工作人员对选课活动刚结束时的筛选表进行监视, 如果发现总选中人数 T 与预置上限人数 M 相差较多时, 则对初始选课概率做出调整: $T > M$, 则下调初始概率; $T < M$, 则上调初始概率。上调和下调的幅度由教务工作人员掌握。

4 应用案例

采用该算法的网络选课系统已经过两次选课过程, 实践证明, 该算法的采用使得选课系统性能更加优化, 主要体现在以下几个方面:

(1) 可承担的并发访问人数增加

图 2 是使用 Nagios 监测系统对选课一周的请求人数进行监测的结果, 可以看出选课系统可同时负担 1500 人以上的并发访问, 这个数目高于以前未采用

该算法时的选课系统。

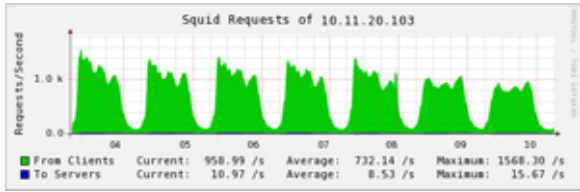


图 2 一周请求人数图

(2) 体现了选课活动的公平性原则

图 3 是对选课一周的网络流量进行监测的结果，结果表明学生不是集中在选课初期，而是均匀的分布在选课时期进行选课活动。由于该算法的采用，使得每个学生具有几乎均等的选课概率，体现了选课活动的公平性原则。

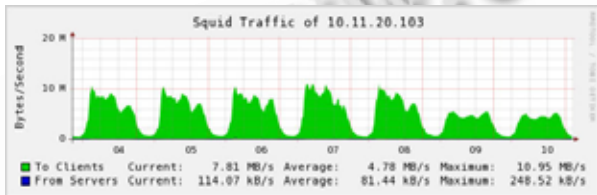


图 3 一周网络流量图

(3) 极大地缓解系统压力、减轻负载

图 4 是对选课全过程的负载情况进行的监测结果，对比未采用该算法时的负载情况(如图 5 所示)，可以看出选课初期负载过重的情况已经没有，该算法的采用极大地缓解了负载过重对系统造成的压力。

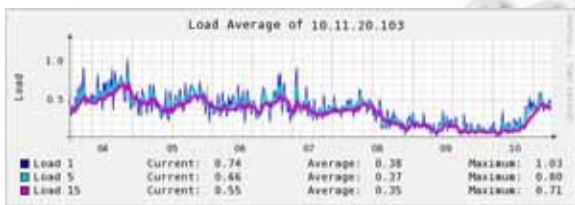


图 4 采用算法后的负载情况图

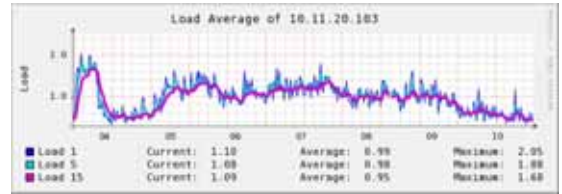


图 5 未采用算法时的负载情况图

5 结语

选课系统的关键是如何解决选课活动的公平性问题及大批量学生并发访问系统带来的网络拥堵问题。本文设计的基于变化概率的公平选课算法，使得每个学生选课活动中拥有趋于均等的选课概率，从而使学生选中与否和其选课顺序无关，既极大的缓解了网络拥堵问题，又解决了选课活动的公平性问题。

参考文献

- 1 蔡志文,万力勇,杨俊峰.基于 Internet 的网上选课系统的设计与实现.成都信息工程学院学报, 2004:25 - 31.
- 2 李冰颖.学分制模式下网上选课系统的算法探析.江西科学, 2004,22(5):358 - 360.
- 3 王帮海.基于分散校区网上限时选课系统的研究与实现.计算机应用研究, 2005:232 - 233.
- 4 关慧.网上选课系统的设计与实现.沈阳化工学院学报, 2004,18(4):295 - 298.
- 5 唐勇.基于遗传算法的排课系统.计算机应用, 2002, 22(10):93 - 94.
- 6 李振坤,孙延海.基于知识库的学生选课系统的设计与实践.计算机应用研究, 2005:721 - 722.
- 7 车战斌,贾晓辉.学分制网络化选课系统及其应用.郑州纺织工学院学报,1995:21 - 24.
- 8 宣华,郭大勇,顾佩.学分制下学生选课管理的研究与实践.清华大学教育研究, 2005,(11):52 - 53.
- 9 郭大勇,宣华,邓伟.清华大学实验教学管理模式改革的新探索—实验课选课系统的设计与实现.高等理科教育, 2005,(6):24 - 25.