

文章编号: 1005-8451 (2013) 10-0017-05

# Erlang模型在坐席规模预测中的研究与改进

李先进, 于展琦, 张仕俊, 张国平

(中铁快运股份有限公司, 北京 100070)

**摘要:** 坐席规模预测是呼叫中心运营管理中的一项重要课题。本文首先对相关排队论模型进行了分析比较, 运用更符合实际话务特征的Erlang A模型对某铁路企业呼叫中心坐席进行预测, 给出了详尽的计算过程。在模型参数选择过程中详细分析了大量历史数据, 充分考虑了企业业务特点, 预测结果符合目标绩效指标和企业实际工作情况。

**关键词:** 呼叫中心; Erlang A; Erlang C; 坐席预测; 4callcenter

**中图分类号:** U293.22 : TP39 **文献标识码:** A

## Research and improvement on agent forecasting using Erlang models

LI Xianjin, YU Zhanqi, ZHANG Shijun, ZHANG Guoping

(China Railway Express Co., Ltd., Beijing 100070, China)

**Abstract:** Agent forecasting was very important in management of operations call center. In this paper, the relevant queuing model was analyzed and compared firstly, then using Erlang A model to predict call center agent with calculation method in detail. The paper analyzed a large amount of historical data and took fully consideration of business characteristics in the choice of model parameters. The predict results could meet the target performance level and conform with work reality.

**Key words:** call center; Erlang A; Erlang C; agent forecast; 4callcenter

在美国, 占全国 4% 的劳动人口工作在超过 20 万个呼叫中心中, 人力成本占呼叫中心运营成本的 70%<sup>[1]</sup>。如何在保证客户服务水平的前提下实现运营成本的最小化和利润的最大化是铁路企业运营管理呼叫中心的一项重要课题。

本文应用排队论模型, 结合某铁路运输企业呼叫中心实际数据对该呼叫中心坐席规模进行预测。模型参数选择过程中详细分析了大量历史数据, 充分考虑了企业业务特点, 结果分析表明预测结果符合目标绩效指标和企业的实际业务特点。

### 1 坐席预测模型选择

呼叫中心的呼叫流程如图 1 所示, 客户的呼入电话在进入呼叫中心后会首先进入长度为 K 的呼入队列, 再由自动呼叫分配器 (ACD, Automatic Call Distributor), 将呼叫分配给 N 个坐席, 完成相关业务。对于有 K+N 条话路的呼叫中心, 当并发呼叫量大于 K+N 时会产生呼损, 客

户无法呼入到呼叫中心, 部分客户可能再拨, 也可能直接放弃呼叫; 当并发呼叫量大于 N 小于 N+K 时, 客户呼叫进入队列等待空闲坐席服务, 客户在被服务前如果失去耐心就会放弃呼叫, 部分可能重拨, 也可能彻底放弃<sup>[2]</sup>。

大部分呼叫中心中继线路充足排队容量较大, 一般不会发生呼损现象, 所以进入 ACD 后的建模是呼叫中心模型研究的重点。此部分的两个重要模型便是 Erlang C 模型和 Erlang A 模型。

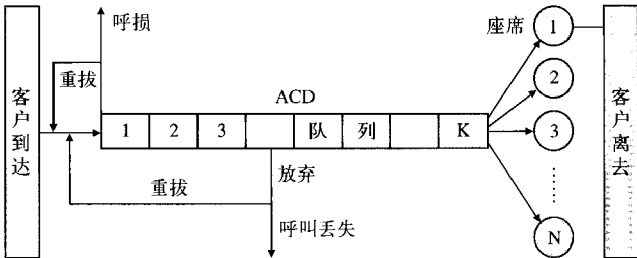


图1 呼叫中心示意图

#### 1.1 Erlang C模型

Erlang C 模型即 M/M/N 排队模型, 是呼叫中心排队模型中最简单、使用最频繁的模型。该模型描述了在话务到达率、平均服务时间、N 个并行服务坐席均为给定条件时, 理论上 N 个坐席都

收稿日期: 2013-02-26

作者简介: 李先进, 高级工程师; 于展琦, 工程师。

被占用或者客户在接受服务之前必须排队的概率。

Erlang C 模型的成立是建立条件：(1) 用户的特征以及坐席的服务技能在统计意义上没有差异，且不随时间变化；(2) 服务顺序为先到先服务无优先级；(3) 不考虑客户到达过程和服务过程中出现的放弃、回拨等行为。

在话务到达时间间隔服从参数为  $\lambda$  的泊松分布， $N$  个坐席服务时间相互独立，服从参数为  $\mu$  的指数分布，其状态转移图如图 2(a) 所示。

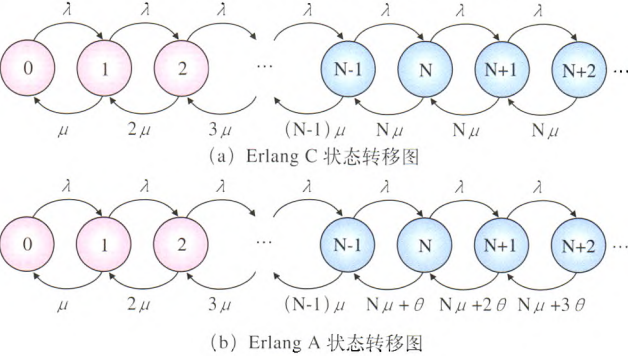


图2 状态转移图

由状态转移图可得局部平衡方程如下：

$$\begin{cases} n\mu P_n = \lambda P_{n-1}, 0 \leq n < N \\ N\mu P_n = \lambda P_{n-1}, n \geq N \end{cases}$$

设  $\rho = \frac{\lambda}{N\mu}$ ，又知  $\sum p_n = 1$  求解可得：

$$P_n = \begin{cases} \frac{N^n}{n!} \rho^n P_0, 0 \leq n < N \\ \frac{N^N}{N!} \rho^n P_0, n \geq N \end{cases}$$

$$P_0 = \left[ \sum_{n=0}^{N-1} \frac{N^n}{n!} \rho^n + \frac{N^N}{N!} \frac{\rho^N}{1-\rho} \right]^{-1}$$

Erlang C 公式的优点：能够将在随机环境下短时间由于系统能力不足而导致的排队行为进行量化，可以为运营管理者提供重要的绩效指标参数。缺点：Erlang C 模型没有考虑呼叫的丢失、客户的放弃以及客户放弃后又重拨等因素。

1.2 Erlang A模型

Erlang A<sup>[3]</sup> 模型即 M/M/N+M 排队模型，该模型对 Erlang C 模型进行了完善，增加了对客户放弃行为的考量。在模型参数：在 Erlang C 模型所需要参数的基础上增加了顾客平均耐心等待时间，并服从指数分布，其状态转移图如图 2(b) 所示，由状态转移图可得局部平衡方程如下：

$$\begin{cases} n\mu P_n = \lambda P_{n-1}, 0 \leq n < N \\ \lambda P_n = (n\mu + (n-N)\theta) P_{n-1}, n \geq N \end{cases}$$

求解平衡方程可得：

$$P_n = \begin{cases} \frac{(\lambda/\mu)^n}{n!} P_0, 0 \leq n < N \\ \prod_{k=N+1}^n \left( \frac{\lambda}{n\mu + (k-N)\theta} \right) \frac{(\lambda/\mu)^N}{N!} P_0, n \geq N+1 \end{cases}$$
$$P_0 = \left[ \sum_{n=0}^N \frac{(\lambda/\mu)^n}{n!} + \sum_{n=N+1}^{\infty} \prod_{k=N+1}^n \left( \frac{\lambda}{n\mu + (k-N)\theta} \right) \frac{(\lambda/\mu)^N}{N!} \right]^{-1}$$

1.3 Erlang C与Erlang A比较

表1 Erlang C和Erlang A重负载性能对比表

	Erlang C	Erlang A	差值	
平均应答时间	68.95	3.23	65.71	↓
平均队列长度	22.22	1.13	21.08	↓
坐席占用率	96.67%	91.00%	5.66%	↓
放弃率	0	5.86%	5.86%	↑

40 坐席，每 15 min 290 个呼叫，平均处理时间 2 min，  
平均耐心时间 1 min

通过表 1 可以发现，其它参数相同仅增加 1 min 平均耐心时间的情况下，同 Erlang C 相比，Erlang A 的平均应答时间和平均队列长度下降近 20 倍，产生 5.86% 的放弃率，坐席占用率下降 5.66%。与表 1 相同呼叫到达率，平均处理时间和平均耐心时间，不同坐席下，平均应答时间、坐席占用率与坐席数量关系如图 3 所示。

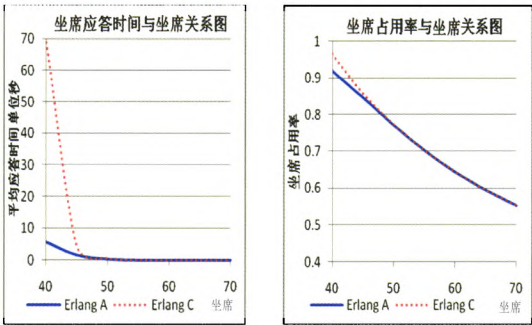


图3 Erlang C和Erlang A对比图

通过对比可以发现当负载较高时，Erlang C 的性能会急剧下降，相同呼叫到达率下需要增加 4 ~ 5 个坐席才能获得与 Erlang A 相同的参数指标；当系统负载较低时 Erlang A 和 Erlang C 性能参数趋同。由于实际运营中确实存在用户放弃的情况，且在系统负载较重时会产生较大误差，所以考虑用户放弃行为的 Erlang A 能够更为准确的



反映系统各方面指标，更适合作为坐席预测模型。

2 坐席预测与结果分析

Erlang A 能够更为准确地反应呼叫中心的实际运行情况，本文以该模型对某企业的单技能组 7\*24 呼叫中心进行分析，使用以排队论理论为基础的 4CallCenters<sup>[4]</sup> 软件进行坐席人员预测和结果分析。4CallCenters 软件可基于多种排队模型进行仿真计算，本文主要应用其 Advanced Profiling 和 Advanced Queries 两个工具在配置为 Erlang A 模型下对相关数据进行分析仿真。

坐席预测 IDEF0 结构图如图 4 所示。

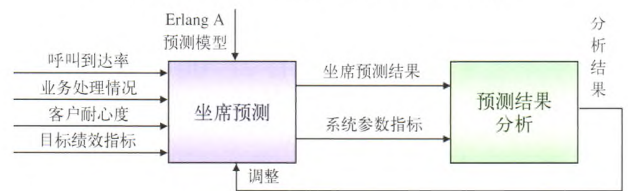


图4 坐席预测IDEF0结果图

2.1 模型参数计算

通过研究发现，Erlang A 模型对系统的呼叫到达率、服务水平、坐席数量、业务处理时间相对比较敏感，对于放弃率等其它参数敏感度相对较低<sup>[5]</sup>，所以在坐席预测中呼叫到达率、业务处理时间和服务水平是模型输入的重要参数。

2.1.1 呼叫到达率

表2 周模型话务量分布表

	周一	周二	周三	周四	周五	周六	周日
话务量	2 403	2 034	1 987	2 024	2 094	1 664	1 438

呼叫到达率预测采用对数正态拟合法对该企业 17 个月的话务数据进行计算，得出呼叫到达率周模型和日模型，周模型话务量峰值结果如表 2 所示，日模型参见图 7 中的呼叫到达率。

2.1.2 业务处理时间

本铁路企业呼叫中心特点：（1）该呼叫中心业务较为复杂，很多业务在客户呼入时不能当场解决，需要坐席人员通过 3 方通话或客挂断电话后联系其他业务部门获得相应结果后再反馈给客户。（2）该呼叫中心每个班次有固定比例的休息时间，具体休息时间由坐席自由随机安排。综上，该呼叫中心的不能只简单计算通话时长作为模型的输入变量，应该综合考虑呼入业务的各方面情况计算得出业务处理时间，将该变量作为模型的

输入变量。结合历史数据和业务自身特点，对呼入、呼出、通话后工作时间以日为单位统计每个呼叫的平均时间（单位秒），呼出呼入比以月为单位统计每月平均比例系数，并定义业务处理时间如下：

业务处理时间 = (日均呼入时间 + 日均通话后处理时间 + 呼出呼入比 \* (日均呼出时间 + 日均通话后处理时间)) / 坐席使用率

该公司 1 月 ~ 7 月上旬日均呼入时间、呼出时间和通话后工作时间数据如图 5 所示。

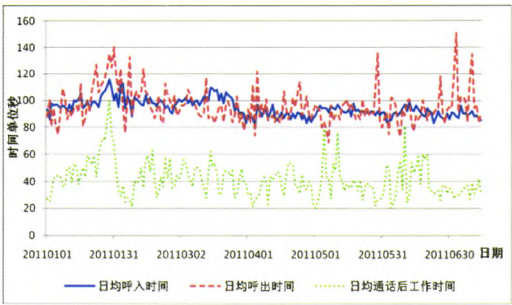


图5 日均呼入、呼出、通话后工作时间时间序列图

可以发现，日均呼入时间较为平稳、日均呼出时间和日均通话后工作时间波动较大，无明显周期性规律，且两者在春节前后由于业务变化较大，同时存在较大波动，其它时段存在部分异常点。

由于该呼叫中心的呼入呼出业务以月为单位进行处理，所以以月为单位进行呼出呼入比计算，15 个月呼出呼入系数如图 6 所示。

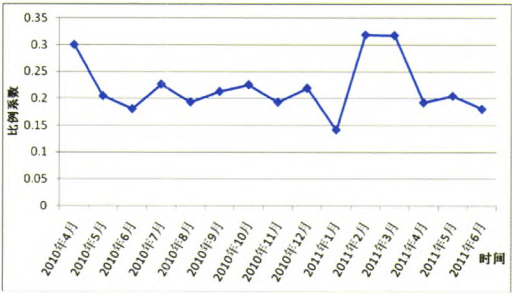


图6 呼出呼入比时间序列图

采用抽样统计的方式对呼叫中心通话处理时间和其它相关参数进行分析，剔除春节前后波动较大数据和其他异常点。数据整理后发现日均呼入时间、日均呼出时间、日均通话后工作时间和呼出呼入比围各自绕某一常数在一定范围上下浮动并趋于平稳，符合呼叫中心业务规律，故采用均值法进行估算。

呼叫中心规定每小时休息 10 min，休息时间



可在整个班次工作时间内累计随机分配，如白班工作 8 h 可随机分配休息 80 min。由于缺少该部分历史数据，本文使用常数进行估算，各参数计算结果如表 3 所示。

表3 业务处理时间计算表

日均呼入 时间	日均呼出 时间	日均通话后 工作时间	呼出 呼入比	坐席 使用率	业务处理 时间
93.351	93.744	39.940	0.203	0.833	192.592

2.1.3 其他参数

客户耐心度，由于该呼叫中心历史数据较少，取某月平均值 40 s。目标绩效指标指呼叫中心运营管理时制定的话务和服务指标，话务指标主要包括服务水平、平均应答时间、接通率等。本文结合该企业需求设定了接通率和服务水平相结合的目标绩效指标，90% 接通率和 20 s 的平均应答时间 85% 服务水平。

2.2 坐席预测

将相关参数带入 Erlang A 模型进行上下限仿真计算，周一坐席预测结果和性能指标如图 7 和图 8 所示。

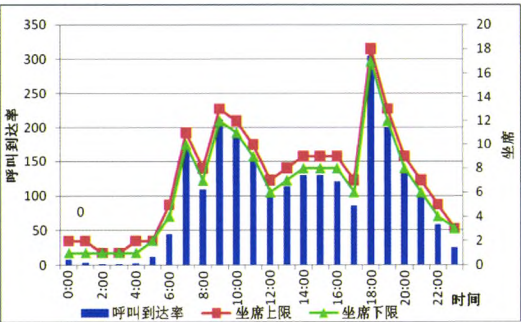


图7 周一预测坐席与呼叫到达率关系时间序列图

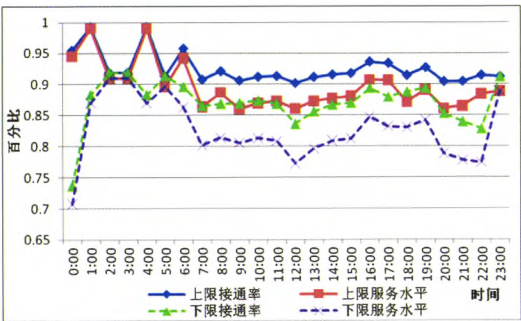


图8 周一预测结果服务水平、接通率时间序列图

预测坐席的上下限值差距一般在 1 个坐席，但坐席人数的轻微变化会导致接通率和服务水平

的较大波动，导致部分区段无法满足系统目标，并且当接通率保持在 90% 时，服务水平会自然保持在 85% 附近。

当呼叫中心坐席不充足或以经济效益为中心时应以下限值做为坐席预测结果；如坐席充足并以服务质量为中心应优先选择系统上限值做为坐席预测结果；在兼顾效益和服务质量的情况下，结合黄金时段黄金观感原理<sup>[6]</sup>将日均服务水平和接通率控制在目标附近。采用最后一种方法对坐席进行预测，具体原则如下：(1) 7:00、9:00、10:00、11:00、14:00、15:00、18:00、19:00、20:00，以上话务高峰时段采用预测上限值，保证系统较高服务水平；(2) 8:00、12:00、13:00、16:00、17:00，以上各时段如接通率和服务水平不低于目标 5%，当日 21:00 ~ 次日 6:00 不低于目标 10% 选用系统下限作为预测值。

2.3 预测结果分析

采用上下限预测兼顾效益和服务质量策略下，周一~周日预测结果如图 9 所示。通过该图发现坐席需求随时间变化呈现明显波动，从日模型看，每天在 7:00, 9:00, 18:00 出现 3 个大高峰时段，最多需要 18 个坐席同时在线，8:00 和 12:00 ~ 17:00 坐席需求量较为平稳，需要 5 ~ 9 个坐席同时在线，夜间坐席需求量较低，仅需 1 ~ 2 个坐席即可。从周模型上看，周一需要坐席数量最多，周二~周五需求数量次之，周六日需求量最少。综上，坐席预测结果符合话务规律。

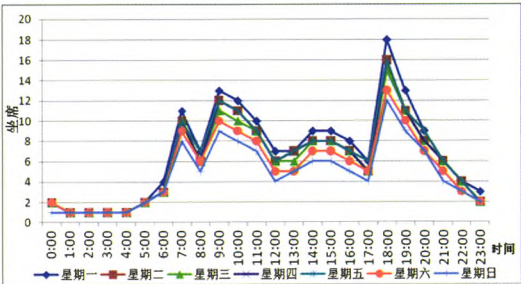


图9 周一至周日坐席人数时间序列图

由图 10 可以知，预测结果的日均接通率稳定在 90% 附近，服务水平稳定在 85% 以上，达到了坐席预测目标绩效指标。日均坐席占用率稳定在 50% 左右，符合坐席实际工作负载情况，且保证了一定的安全冗余。

(下转 P23)



车次按照临时、图定、专运进行分类；右侧针对具体客车车次交路进行维护。图3展示了T62\T61车次的交路信息：(1) 客车信息包括开行区段、运行周期、图定属性；(2) 交路图包含客行走行车次和时刻表；T62\T61有4组车，每组车从昆明始发时间，这个时间由系统自动生成；(3) 车组车底配置要求：T62\T61每组车有16个车厢，1位~16位所配备的车型和定员要求，为编制具体车次车组提供模板。

本系统中交路图维护的客车车组的开行日期是通过开行周期自动推算出来，如T62的1组系统确认已在8月5日开出，将根据4天的周期，推算出下次开出时间为在8月9日。这个计算是通过Oracle定时任务完成的。如果出现因特殊情况客车停开，维护人员可手动对客车开行时间进行调整，使系统兼顾了自动和灵活特点。

### 3.3 车辆信息库维护

该功能主要是车辆调度和客车车辆段使用。实现对全局车辆库的维护。包括增加、删除、修改等。

(上接 P20)

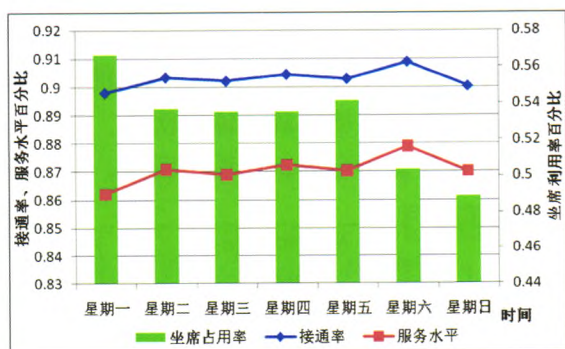


图10 预测结果的日均坐席占用率、服务水平和接通率时间序列图

## 4 结束语

本文首先分析了Erlang C和Erlang A模型的原理与区别，采用能更为准确反映呼叫中心话务特性的Erlang A模型对某呼叫中心坐席进行了预测，扩展黄金时段黄金观感原理解决预测中的结果和参数选择问题，给出了详细的参数计算方法和过程，预测结果很好的匹配了目标性能。由于缺少相关历史数据，本文没有对放弃率进行详细分析。考虑到铁路行业的自身业务的特殊性，对

## 4 结束语

铁路客运车辆调度信息系统自2012年在昆明铁路局投入使用以来，通过采集车辆段运用客车的基础信息数据，拣牌操作便捷高效，丰富了车辆管理信息内容，实现了在途列车昆明出发日期的自动跳转，能够对人员操作进行审计，有利于事后的责任划分和统计分析，能够自动发现与编组计划不符的有关事项，并报警提示，进一步减轻了工作量，节约制作车号牌的资源，适用性和准确性强，实现了原定的系统设计目标和功能，目前，正在研究利用GIS技术反映开行客车信息的方法。

### 参考文献：

- [1] (美) Ying Bai. C#数据库编程实战经典[M]. 施宏斌, 译. 北京: 清华大学出版社, 2011, 11.
- [2] (美) Michael McLanughlin. Oracle Database 11g PL/SQL 程序设计[M]. 张云, 王海涛, 译. 北京: 清华大学出版社, 2009, 4.

责任编辑 徐侃春

呼叫中心复杂业务的运营方式也需要进一步研究。

### 参考文献：

- [1] Mandelbaum A. Quality and efficiency driven queues (with a focus on call/contact centers)[J]. Euro Working Group on Stochastic Modeling, Koc University, 2008, June:23-25.
- [2] Gans N, Koole G, and Mandelbaum A. Telephone call centers: a tutorial and literature review[J]. Invited review paper, Manufacturing and Service Operations Management, 2003, 5(2):79-141.
- [3] Palm C. Research on telephone traffic carried by full availability groups[J]. Tele, 1957, 1:101-107.
- [4] 4CallCenters Personal Optimization Tools for Call Centers [EB/OL]. <http://iew3.technion.ac.il/serveng/4CallCenters/Downloads.htm>.
- [5] Ward Whitt. Sensitivity of Performance in the Erlang A Model to Changes in the Model Parameters[J]. Operations Research, 2006, 54:247-260.
- [6] 许乃威. 两“困”相权, 孰为轻, 抉择与决策[EB/OL]. <http://www.51callcenter.com/newsinfo/154/10032/>, 2007-2-27.

责任编辑 徐侃春