Incident Ticket Analytics for IT Application Management Services

IT 应用管理服务的事件票分析

Ta-Hsin Li, Rong Liu, Noi Sukaviriya, Ying Li, Jeaha Yang, Michael Sandin, and Juhnyoung Lee IBM T.J. Watson Research Center

Yorktown Heights, NY 10598-0218, USA

e-mail: {thl,rliu,noi,yingli,jeaha,msandin,jyl}@us.ibm.com

摘要

一个重要的 IT 服务外包业务是解决与我们公司支持的客户合同相关的 IT 基础设施事件。事件被记录作为门票结构化和非结构化数据,其中包含各种特征的事件包括时间戳、描述和解决。分析此类事件票成了一个关键的任务在管理操作服务用来保证在约定服务水平协议的操作。工作票分析是至关重要的识别异常情况和趋势，以及检测不寻常的操作模式，这种分析是很难做到手动的尤其是大客户复杂的组织和范围。本文主要关注工作票分析和一些关键统计技术应用于分析。最后，我们使用真实数据示例演示这些技术，并讨论工作票分析的主要挑战。

关键字

事件管理，IT 服务管理，票务分析

1.介绍

AMS 是一种使用定义良好的 IT 外包服务，全球一体化的进程，

政策，程序和标准来管理客户的应用程序组合。帮助台的提供通常包括服务，应用程序的维护和支持和应用程序健康监测。对于大多数企业来说，一个应用程序可能是一个软件产品如 SAP 和 Oracle 模块，或解决方案如ＣＲＭ或特定的软件。事件票是记录服务请求处理失败，错误，或任何问题与应用程序支持的合同。它包含结构化和非结构化数据的结合。结构化数据的例子在票请求类型，应用领域，该组织的轨迹问题，解决票的请求的时间戳和决议，在某些情况下，票解析器的主要时间花在解决票上。非结构化数据的例子有文字描述问题和文档的决议。票的广泛目标分析包括：

1. 票卷的指示操作的评估工作量；

2. 评估票决议时的操作效率；

3. 识别潜在的事件造成的问题。

我们已经开发了一个基于 web 的工具【9】提供票务分析作为 AMS 客户的标准服务。该工具需要操作数据的客户端和其他结构化数据票作为输入，自动计算指标，评估这些指标，并提供一个仪表板总结客户的操作性能。虽然【9】提供了详细的工具，本文将重点分析和深入阅读核心票使用的统计方法分析。我们还将提供一个案例研究演示一些技术和他们的商业价值。

本文的其余部分组织如下。第二节提供了一个真实的业务用例背景下票的分析，基于我们的一个客户端项目。第三节给出了一个简短描述的事件票数据结构和 ticket-related 指标。第四节详细描述了统计技术以及它们的业务价值。第五节显示了基于 web 的票分析工具

的体系结构。第六节比较这项工作与相关工作。最后第七节总结本文简要描述我们的未来的工作。

2.为什么分析票—一个业务案例

基于 web 的票分析工具【9】已经被几十个全球公司的各部门和行业在几年内协助 AMS 操作。在本节中，我们描述一个代表性的使用这个工具通过展示典型票分析是针对解决的业务问题,及其在帮助提高 AMS 的操作性能上的商业价值。

客票数据的分析方法被用于客户端项目回答一些业务问题事件管理操作。下面是一些典型问题的关键：

1. 比较的问题
   a. 有没有票数量在不同的应用程序或组织存在显著差异?
   b. 实际的票体积是如何分布在优先级或严格的符合最初的假设？
   c. 在解决不同时代的票的类别或团体上是否存在显著差异？
   d. 票决议是如何满足服务水平协议（SLA）的要求的？
2. 热门问题
   a. 票量随时间变化如何？有任何重要的模式或趋势吗？
   b. 票决议时间随时间变化如何？
3. 预测问题
   a. 下一个时间段（月，季度，等等）预测的是什么票卷？
4. 票技术结构和解决策略的问题

a. 哪些典型问题症状或有主要问题的地区导致了大多数的票？

b. 有哪些常见的解决不同类别的票吗？

c. 有可能自动票决议吗？

这些问题的答案可以极大地帮助团队在他们的日常操作中做决定。例如，知道机票跨应用程序的负载变化可以帮助配置具有适当技能的人员。这是更重要的是当资源跨应用程序不共享是由于不同技能要求的应用程序。热门问题可以帮助识别表现不佳的业务领域。例如，趋势解析时间在票卷或应用程序上没有显著变化可能表明有辱人格的受托人的性能或效率低下的员工。

3.事件票和指标

事件票是一个重要的工具测量和 AMS 质量实现。票分析利用原始结构化数据和/或有时从原始票数据计算得到的 custom-specific 指标。例如，操作工作量通常表示为一组可衡量的指标包括票卷（每月，每周，每天），每个应用程序或其他类别，待办事物列表的票量（收到的票的数量和在特定时间内及时完成的门票的数量）。这对模型票数据是很重要的以至于有足够的信息机制提供操作的见解。在本节中，我们提供一些详细的数据结构作为后续分析的背景部分。

图 1 显示了一个示例的机票，它包含以下关键属性：

• 票ID：案件的唯一标识符 ；

• 票务开放时间：时间戳案例提交时显示；

• 票解决时间：时间戳指示此案时解决；

• 票关闭时间：时间戳指示案件时关闭；

• 票状态：票据的处理现状；

• 票决议时间：票关闭/解放时间和票开放时间之间的消逝可以在日历天/小时，或业务天/小时。具体来说，在后一种情况下，我们需要作出某些假设票解析器的营业时间。例如，他或她每天工作8小时，每周五天。然而，这样的定义可以随票的优先级而不同。对于票决议时间的更精确测量，额外的元数据描述需要票解析器的营业时间。

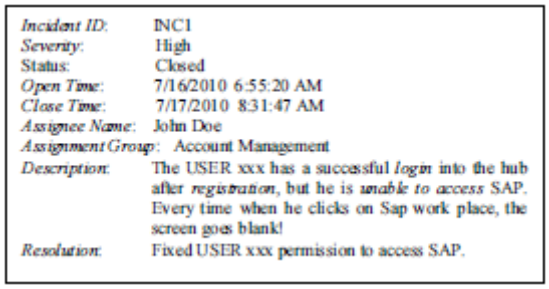
Incident ID: INC1
Severity: High
Status: Closed
Open Time: 7/16/2010 6:55:20 AM
Close Time: 7/17/2010 8:31:47 AM
Assignee Name: John Doe
Assignment Group: Account Management
Description: The USER xxx has a successful *login* into the hub after *registration*, but he is *unable to access* SAP. Every time when he clicks on Sap work place, the screen goes blank!
Resolution: Fixed USER xxx permission to access SAP.

图1　带有典型属性的样本事件票

此外，票数据包含上下文属性，支持分组分析的门票。这些属性如下：

• 票决议代码：分类的解决方法；

• 票受让人（或任务组）：名称或ID的技术人员（或一组）分配给解决车票；

• 应用程序：名称创建应用程序导致的票；

• 票严重性或优先级：严重程度的分类或优先级的票。

在一些票系统中，票开和票关闭之间的时间进一步划分为一些额外的部分，如分配时间（时间戳表示当被分配到一个技术人员）和解决时间（时间戳表示此案例被技术人员解决但在等待客户的批准关闭）。此外，票决议代码或应用程序参与门票经常没有明确记录在票，由于缺乏明确的分类模式。在这种情况下，文本挖掘的票文本可以替代基于票的方式组症状或解决策略。

基于事件的门票,可以计算“指标”。例如,票量表示收到的门票在一段时间内， 音量可以进一步按属性分组,如严重程度,应用程序,任务组等。另外一个例子,决议时间可以聚合和平均通过时间和/或分组属性。

指标计算后,我们应用一个票操作评估过程监控这些指标,分析他们理解的正常行为,识别异常情况,并提供可操作的改进的建议。这个过程不仅需要巨大的领域知识,而且还有许多统计方法，例如,统计过程控制(SPC)[11]对于异常检测和文本挖掘识别主要问题的原因记录在文本的机票。

在一个稳定的商业环境,这些指标通常不随时间急剧波动。然而,这些指标的不寻常的运动可能反映了重大的商业环境的变化,例如,

新应用程序发布,员工流失率,和应用程序环境问题。因此,分析这些指标时间序列可以披露深层原因和建议补救措施的性能改进。

4.统计方法和机票的例子分析

如前所述,使用统计方法来分析结构化和非结构化数据事件票。在本节中,我们将详细描述一些重要统计方法应用于我们的分析。4.1到4.4节将关注那些应用于结构化数据在4.5节描述数据挖掘数据结构化的票。

A. 票卷分析统计过程控制

统计过程控制(SPC)在制造业监测质量的部分和检测生产过程中的潜在缺陷是一种广泛使用的技术。它也采用服务行业服务质量控制的工具。在这项工作中,我们应用SPC方法监控每周/每月,积压的门票,和每周/每月决议时间包括标准偏差和四分位数统计信息。有很多方法可以建立一个程控程序。

一个简单的例子是一个图表,描述了变量的观测值的序列,我们一起要监控一个或两个水平线条,描绘3-sigma规则,即。,3个标准差(σ)上方和下方的平均(后者可以省略,如果异常低的变量的值是不关注的用户)。任何数据点交叉3-sigma行意味着平均偏差超过3倍的标准差,可以标记为异常(异常值)。与生产应用系统漂移表示在制造过程中的潜在问题不同,机票的销售量的趋势预计是由于业务的扩张与伸缩。季节性机票销售量预计服务业务时,周期性的行为(例如,零售行业)。在这样的情况下,异常检测应由考虑到预期的趋势和季节

性,和标准程控程序删除后可以应用趋势和季节性。这一趋势和季节性模式可以通过最小二乘回归估计明确技术。这些特性是重要的容量规划的目的。图 2 显示了一个例子,趋势和季节性的解除对异常检测产生了影响。

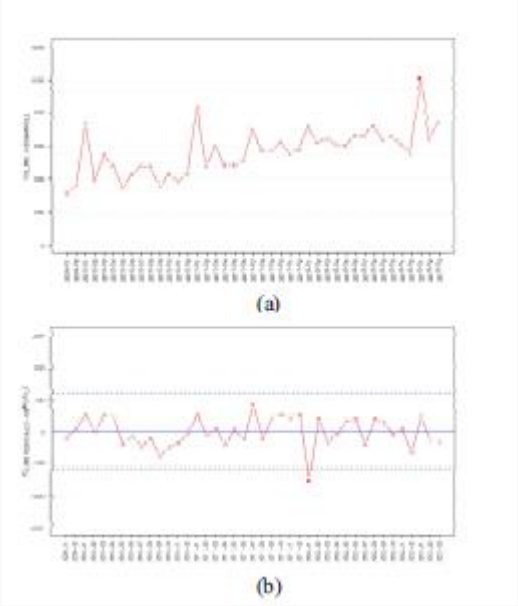

图 2 监控月票卷由一个控制图与上、下极限定义的平均值加减 3 倍标准偏差而来。(a)原始时间序列。(b)切除后剩余时间序列趋势和季节性。在这里,实线表示平均数,而虚线表示上限和下限。

图 2（a）从一个实际的客户端描绘了每月的某一票类型。该系列显示出强劲的增长趋势与偶尔的峰值和其他更复杂的变化。在 2013 年 1 月，一个简单的应用程序控制图的上限和下限的平均值加减 3 倍标准偏差发现轻微异常值与异常高容量，从业务的角度令人担忧。

然而，在前几年通过仔细检查卷，我们可以看到，一月往往是一个月的量尖峰,但这些峰值不超过上限,因为他们的增加趋势是不明

的。在此背景下检测异常趋势和季节性,我们首先假设是线性趋势和季节性自然是 12 个月。由此产生的模型可以表示为

$$m(t)=a+bt+\sum ck\ I(t=k\bmod 12)\ k=1\quad (t=1,2,...).$$

这个模型的系数是由最小二乘估计。通过删除原始数据 v（t）的估计趋势和季节性,我们获得剩余时间系列 r(t)= v(t)- m(t)。图 2(b)与图 2(a)以同样的方式显示剩余时间序列中定义的控制限度。事实证明,把趋势和季节性考虑在内的话，2013 年 1 月时不再是极限值。相反,控制图表显示,2012 年 1 月有一个异乎寻常的低容量，从企业的角度，可能一点也不令人担忧。

另一方面,监控积压的体积通常是比普通票卷低的,所谓的 c-chart 似乎更合适。图 3 显示了一个真正数据的例子。这个图表类似于图 2 所示的除了上限被定义为平均值加 3 倍的平方根的平均而不是标准偏差的 3 倍。c-chart 更适合监测小数量的时间序列，底层常态假设在图 2 中变得不那么必要。c-chart 基于泊松假设的下限低于 0,而且有一个更长的上尾成正态分布。

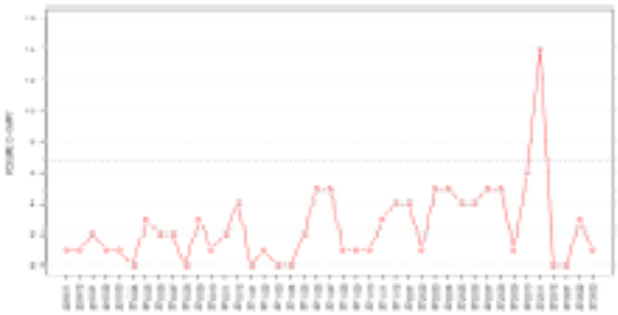

图 3 监控每月由 c-chart 积压的上限定义为平均值加平均值 3 倍的平方根

B. 统计过程控制方法对于票决议时间分析

票与票之间的决议时间显著不同。变化取决于机票的性质，以及解决机票的技术人员的技能和性能。要监测票的决议时间的行为，就必须把重点放在一定的有意义的统计上。平均解决时间是这样的统计数字之一。图 4 显示了一个监控的例子，每月的平均分辨率为一个真实的应用程序的时间。虽然图 4 中的控制图（a）检测到 2011 中的异常值，但更近的观察表明，平均分辨率时间经历了一个向上的转变，持续数月。检测这种温和但持续的变化，可以用所谓的 CUSUM 控制图，如图 4（b)。（标准化）的一系列 cusums x（t）的定义为 $S(0)=0$, $S(t)=\max\{0, S(t-1)+[X(t)-\mu-\delta/2]/\sigma\}$ $(t=1,2,K)$, 其中 μ 是参考价值意味着（$\mu=1$), δ 是目标的转变意味着检测（$\delta=0.3$), σ 是标准偏差（$\sigma=0.17$)。其中 μ 是参考价值意味着（$\mu=1$), δ 是目标的转变意味着检测（$\delta=0.3$), σ 是标准偏差（$\sigma=0.17$)。一次向上转移声明 S(t)超过上限通常设定为 4 或 5 h。图 4(b)显示了 cusum 图表与 h = 4。正如我们所见,它检测到一个向上转移期间的 2011 年 7 月到 2012 年 7 月。平均分辨率时间低于检测限后，在 2012 年 6 月，指示改善过的处理的票。

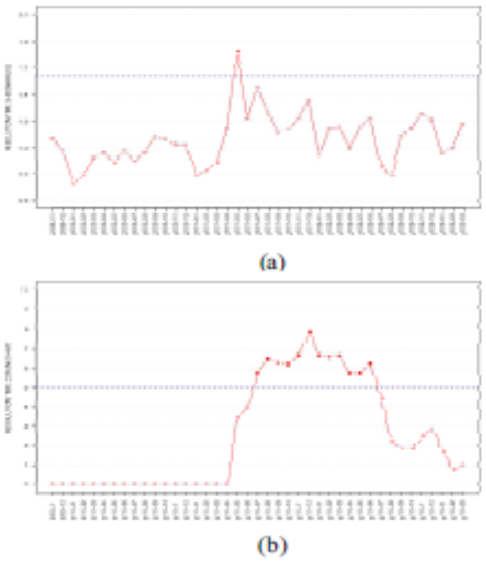

图 4　监测月平均解决时间。（a）控制图的上限为平均加上 3 倍的标准偏差。(b)Cusum 图表上限设定为 5。(a)在这里,实线表示平均值,而虚线(a)和(b)表示上限。

C. 服务质量保证的统计过程控制

对于服务质量保证，基准测试，和许多其他原因，比较不同的机票类别的时间分辨率是很有必要的。图五显示一个月内 5 个不同类别的门票的累积概率分布。

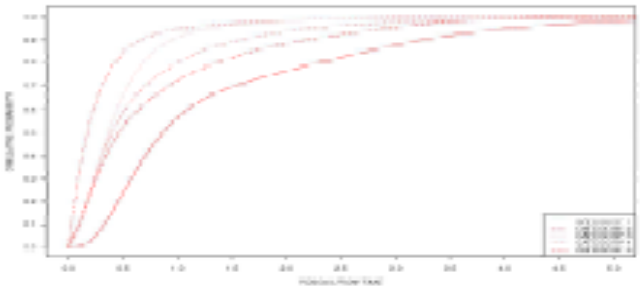

图 5　五类票的概率分布的解析时间

一个简单的参数化模型的票的时间是对数正态分布，其中的分辨率时间的对数变换被假定为一个随机变量的高斯分布，这是索引的参数，均值和方差。在这种假设下，可以进行配对 T 检验比较所有对应用程序的时间分辨率（Holm 1976）。

用yi, si和 ni表示均值，标准差，以及从I类票日志分辨时间的样本大小的一对范畴，i和j，t检验差异均采用标准化的方法，即.

$$t_{ij} = \frac{y_i - y_j}{\sqrt{s_i^2/n_i + s_j^2/n_j}}.$$

如果这个量的绝对值大于预定阈值,然后在他们的决议中i和j的分类被认为是有明显不同的方式。计算出每一对p值比较的统计学意义。这是假阳性的可能性。

Table I contains the t-statistics and their p-values for testing the equality of the means for every pair of the five categories. It shows that eight out of ten pairs have significantly different means, only the 3-4 pair and the 4-5 pair do not have significantly different means.

表1　检验统计量及其五类票分辨时间两两比较P值。整体假阳性率为0.05。

	Pairwise *t*-Test		Pairwise Rank Test	
Category Pair	*Test Statistic*	*p-Value*	*Test Statistic*	*p-Value*
2-1	−30.1	0.00	−0.43	0.00
3-1	−19.8	0.00	−0.34	0.00
4-1	−9.0	0.00	−0.27	0.00
5-1	−8.2	0.00	−0.20	0.00
3-2	20.5	0.00	0.27	0.00
4-2	9.9	0.00	0.26	0.00
5-2	12.9	0.00	0.27	0.00
4-3	1.6	0.21	0.02	0.99
5-3	3.9	0.00	0.06	0.14
5-4	1.4	0.21	0.04	0.75

除了成对t检验，这就需要对日志分辨时间的高斯假设，也可以执行一个基于没有任何假设的分布数据的行列成对比较（konietschke等人。2012）。与成对t检验，秩为基础的测试测量的差异之间的一对分布i和j所谓的相对效应得分定义为

$$d_{ij} = \frac{1}{n_j}\left\{r_i^{(ij)} - \frac{n_i + 1}{2}\right\} - \frac{1}{2},$$

ri(ij)表示该决议的平均排名次x_{ik} (k = 1,...,n_i)在样本类别i和j合并后{ x_{ik},x_{jl}: k = 1,...,n_i; l = 1,...,n_j } 。统计dij是阳性的如果票的解决时间从类别i倾向于超过类别j;它是阴性的如果门票从应用程序j的决议时间超过类别i。如果绝对值统计足够小,那么这两个发行版被认为是几乎相同的。

表1的最后两列包含d-statistics和p值的测试。类测试阳性的七对用配对t检验，发现都是相同的。然而，对组成的类别3和5认为是不可区分的分布。因此，根据秩检验，可以把应用程序3，4和5作为一个集群中的分辨率时间有类似的分布，而类别1和2有不同的分布，以及从集群组成的类别3，4，和5。

D. 票量预测的统计过程控制

最后,让我们考虑票量预测的问题。根据预测卷的行为，有一个广泛的时间序列模型,可用于预测票卷。作为一个例子,考虑月票数量如图2（a）所示。本系列有增加的趋势,以及强大的12个月周期。一个适合的时间序列模型是季节性自回归移动平均（ARIMA）模型，指示为ARIMA(p、d q)×(p、d q)，其中p，d，q代表普通回归，差的订单，和移动平均线，分别，和P，D，Q代表季节性自回归，差的订单，和移动平均。

图6描述了时间序列和12个月的提前预测月票卷图2（a）所示，以庆祝（0,1,1）×（0,1,1）与12个月的周期模型。趋势和季节性都很好地反映在预测中，延长增长的趋势，以及在十二月的强峰和温和

的峰值在七月的季节性。

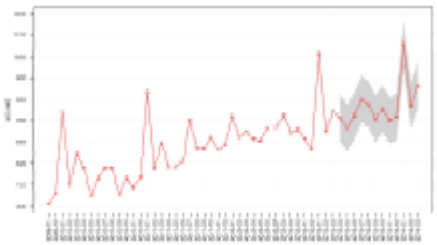

图6　月度预测卷。最后十二个值从2013年4月到2014年3月预测卷，和周围的阴影面积这些值代表了90%置信区间。

E. 聚类门票的症状定位文本挖掘

基于他们的文本描述的集群门票可以提供的见解的技术混合物的门票。有一些有效的聚类技术的工艺文件和文档聚类，例如，K近邻(KNN)，层次聚类，潜在语义分析(LSA)，隐含狄利克雷分配(LDA)，等[1]。作为一个例子，表2显示了门票的聚类结果相关的事务处理应用程序。具体地说，我们以这个应用程序的分类能力首先处理1445张门票相关的文本提取关键字和短语。然后，每张票都被表示为一个向量的提取关键字/短语后的向量空间模型[12]。一对票之间的相似性测量这两个向量之间的相似度。最后，我们采用分层聚类算法的基础上的机票相似的组票。

表2　基于票文本的交易处理应用聚类事件票。结果表明，多数票是由消息队列溢出引起的。

Cluster	Number of Tickets	Percentage

Message queue overflow	864	59.79%
Communication channel errors	190	13.15%
Disk space issues	94	6.51%
Data format mismatch	76	5.26%
Security & user access	58	4.01%
Database issues	22	1.52%
Others	141	9.76%
Total	1445	100.00%

这张票可以对提高聚类AMS操作进一步洞察。如表2所示，任何努力降低票量或自动解决在“消息队列溢出”集群能有效地提高AMS演出的票。此外，该群集创建一个额外的分组属性的门票。我们还可以分析在时间维度上的每一个集群的门票数量或分辨率。

V 机票分析系统构架

图7显示了在一个高层次的网络为基础的机票分析系统的体系结构，显示不同的层，包括数据，工具，操作和演示，以及不同类型的用户的系统带来不同的分析使用场景的系统。系统架构的设计目标是提供一个标准，综合分析平台支持AMS交付采用标准开放协议栈软件在网络平台上建立与先进的分析方法，以提高生产率和增强对AMS实践交付质量。事件的数据通常是专有的系统，记录，并经常以不同的格式存储从一个系统到另一个系统。

而每一层的详细描述在系统中可以找到另一篇论文[9],我们简要讨论工具层。票分析工具报告中描述的分析。目前可以在IBM Cognos商业智能服务器上实现。它让有经验的分析师为用户配置分析报告与

报表参数的有效和有意义的选项和过滤器。不那么复杂的用户和账户管理人员可以利用预定义的报告和仪表板，只需选择预定义的参数和过滤器。

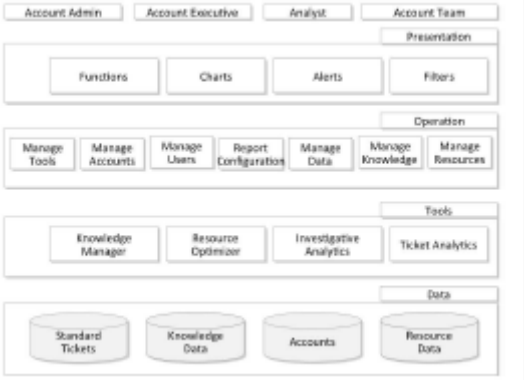

表7 票分析系统架构

调查分析工具可以帮助一个帐户管理和跟踪它的问题根源。调查的过程通常包括几个阶段。该工具提供了各种诊断算法缩小到一个有问题的事件数据的子集。同时,它确保事件诊断过程中连续、无缝体验。

知识管理工具可以使用的事件分辨率从业人员。通过捕获和编辑解决知识经常重复出现的事件和促进知识的从业者在时间约束的环境的重用，该工具有助于提高AMS交付实践。资源管理工具有助于交付管理人员对交付资源的利用，以及在客户帐户和客户帐户中规划资源的需要。

VI 相关工作

IT服务管理（ITSM）是关系到客户的业务对IT贡献的一个新领域。IT服务管理是不同于传统的以技术为中心的方法来管理和业务交互作用[6，13，14]。事件管理是它的一个服务管理过程区[7]。事件

管理的目标是尽快恢复正常的服务操作，并尽量减少对业务运营的影响，从而保证服务质量和保持在可用性的最佳水平。信息技术基础设施库(ITIL)是一组实践ITSM针对于IT服务与业务的需要[4]。事件管理的主要流程包括事件检测和记录,分类和初步支持,调查和诊断、解决和恢复,事件关闭,和事件的所有权、监控、跟踪和沟通。AMS分析系统提出了担忧主要是运营分析、调查和诊断ITSM的事件。

此外,AMS分析系统为知识管理,资源管理和经营预测分析模型提供先进的分析功能。预测分析或预测分析是一个统计分析的领域，从数据中提取信息，并使用它来预测未来的趋势和行为模式[10，2]。预测分析的核心依赖于捕获的解释变量和预测变量之间的关系，并利用过去的事件来预测未来的结果。由于其重要的业务，事件诊断和预测分析中的应用已经讨论通过论坛和其他媒体的从业者[6]。然而，问题是最近解决的一个严格的学术方法。

VII 结束语

在本文中,我们演示了一些票统计技术分析。任何先进的分析方法，这些技术的成功应用需要一定的理解的基本假设和他们的陷阱，特别是当这些技术是自动的机票分析系统。例如，机票卷可能会出现非常复杂的模式，不能充分模拟为一个线性函数加上季节性调整。虽然更复杂的技术可以用来处理异常检测这样的模式，使这些方法的选择套票量的波动范围广泛仍然是一个具有挑战性的任务。没有适当的考虑这些模式，SPC技术可能会产生过多的误报或错过重要的真正的异常。

票积压通常随着时间的推移累积计算，数票开在一段时间内减去门票的数量在这段时间关闭。在这种情况下，他们应该被解释为积累超过时间的门票，而不是积累的门票等待解决。因此，他们是受的企业实践影响关闭一个解决的机票。因此，票都应该在一个频率，与业务实践相兼容的监测。例如，如果机票被关闭后立即决议，开闭幕的时间往往是短暂的（例如，在几个小时）。在这种情况下，应在短时间间隔内（例如，每日）进行积压的计算。另一方面，如果需要几天或几周的时间来接近某一类的请求，那么一周或每月的频率应该是足够的。

票决议时报通常包含异常值,可以很容易地掩盖了计算平均解决时间。因此,更健壮的指标,如中值,应被视为一个可能的选择意味着监测解决时间。

在不同的机票类别的变化，机票的时间可以有很大的不同。这复杂的跨类比较，因为平均分辨率时间不再是有意义的，因为它是当变异仍然是相同的所有类别。在这种情况下，其他指标，如相对影响得分，可以认为是可能的替代品。