

一种 KPI 映射到 KQI 的通用算法

倪萍 廖建新 王纯 张成

(北京邮电大学网络与交换技术国家重点实验室 北京 100876)

(东信北邮信息技术有限公司 北京 100083)

摘要: 该文结合电信管理论坛 TMF 提出的 KPI(Key Performance Indicator), KQI(Key Quality Indicator)概念, 提出了一种新算法 K2K(KPI to KQI)。新算法弥补了 TMF 提出的算法不能聚合不同量纲、不同意义指标为单个 KQI 指标的缺陷, 并且 K2K 算法显著减少了最终聚合指标的个数, 提取少数有意义的指标以直观地反映整个系统的运行状况。实际使用效果的统计数据验证表明, K2K 算法能够真实反映系统的实际运行状况。

关键词: 网络管理; KPI; KQI; 运营数据分析

中图分类号: TP393.07

文献标识码: A

文章编号: 1009-5896(2008)10-2503-04

A General Algorithm Mapping KPI to KQI

Ni Ping Liao Jian-xin Wang Chun Zhang Cheng

(State Key Laboratory of Networking and Switching Technology,

Beijing University of Posts and Telecommunications, Beijing 100876, China)

(EBUPT Information Technology Co., Ltd, Beijing 100083, China)

Abstract: This paper proposed a new algorithm called K2K (KPI to KQI) based on algorithm which had been proposed by TMF (Telemanagement Forum) proposes a general algorithm about how to map the key performance indicator (KPI) to the key quality indicator (KQI). K2K algorithm aggregates kinds of indicators to one single significant indicator to express the whole network status. However the algorithm provided by TMF only aggregates the same unit, so it couldn't extract few key indicators to express the whole network healthy status. K2K algorithm had been implemented in practice. It could express the real running status of systems.

Key words: Network management; KPI; KQI; Operation data analysis;

1 引言

新业务的管理同传统网管相比更注重的是用户的感受, 以及对出现问题的快速察觉能力。面对激烈的竞争, 运营商需要保证业务 QoS(Quality of Service), 提高用户满意度以及自身竞争力。新业务没有现成的量化评价指标来评估其业务 QoS, 而传统的网络性能指标又不能直接准确地反映客户的感知, 网络设备指标好坏与用户感知之间没有明确的映射关系。世界电信管理论坛(TMf)提出了 KPI(Key Performance Indicator)和 KQI(Key Quality Indicator)的概念, 并且给出了相关映射算法^[1], 但是其存在一定的缺陷, 不能聚合不同量纲、不同意义的指标, 如延时和呼损率这两个指标。本文提出的 K2K(KPI to KQI)算法则不会受指标单

位量纲以及指标意义的限制, 并且 K2K 算法大大地减少了聚合指标数, 使得运营维护人员观察整个业务的运行状况更直观, 提高了业务运行质量评价的准确度以及发现、处理问题的效率。本算法已经在实际项目中得到了应用, 取得了良好的效果和较好的经济效益。

2 关键指标层次

对于关键指标的层次结构, 世界电信管理论坛给出了如图 1 所示的架构^[1]。

从图 1 中可以看到, 在最底层定义了底层指标, 是与网

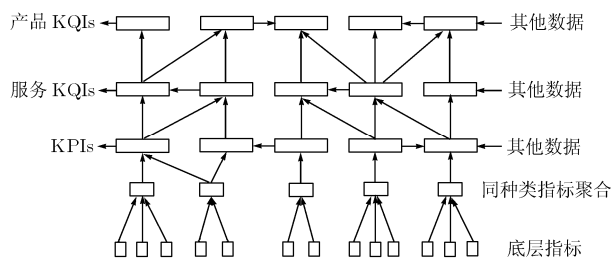


图 1 关键指标层次图

2007-04-12 收到, 2007-09-21 改回

国家杰出青年科学基金(60525110), 国家 973 计划项目(2007CB307100, 2007CB307103), 新世纪优秀人才支持计划(NCET-04-0111), 电子信息产业发展基金项目(基于 3G 的移动业务应用系统), 电子信息产业发展基金重点项目(下一代网络核心业务平台), 电子信息产业发展基金项目(基于内容的综合通信网络计费平台)和国家高技术产业化信息化装备专项项目(支持数据增值业务的移动智能网系统)资助课题

络或者设备性能或者业务运行指标有关的数据。那么该如何映射到上层的服务 KQI, 即得到业务的运行质量? TMF GB917 文档中提到了相关算法^[1]。

TMF 提出的解决方案存在如下缺陷:

(1)不能把不同量纲, 不同意义的指标进行归一化来评价, 例如吞吐量和时延。

(2)不能真正地将底层指标反映到业务层面上, 只是对相同量纲数据的聚合, 结合其不能将不同意义的指标进行聚合必然也会导致下面的问题(3)。

(3)如果要对多个指标进行评价, 比如存在多种量纲如负荷、利用率、点击次数、话务量等, 聚合后会得到各种不同的指标, 不能达到聚合指标尽可能少且直观以便准确评价整个业务运行质量的要求。

本文提出了一种通用的 K2K 算法, 在这 3 方面进行了提高, 具有如下优点:

(1)K2K 算法能把不同的 KPI、KQI 指标进行归一化, 不会受 KPI 单位量纲、指标意义的影响。

(2)对业务运行状况以分数的形式表现其运行质量, 使得客户在业务层面直观感受业务系统运行状况。同时, K2K 算法利用递归的方式由下向上递归得到图 2 层次树中的各个指标所代表的健康状况。

(3)在 K2K 算法模型中, 减少了聚合后的指标数。

3 改进的关键指标层次

改进的关键指标层次树如图 2 所示。

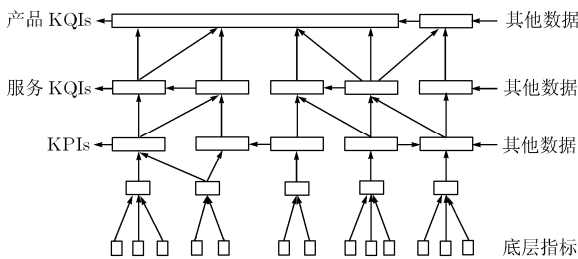


图 2 改进的关键指标层次

底层指标: 对于不同的业务和系统会有不同的指标, 它也属于 KPI, 如图 3 中的中继电路负荷、资源节点磁盘空间等都属于底层指标。

不同种类、不同意义指标聚合: 在聚合底层指标过程中为了增加对系统评价的准确度使得粒度更精确而提出的一种过渡聚合, 比如 CPU 利用率、内存利用率、硬盘可利用空间可以先聚合为某个指标, 例如称其为系统资源利用率, 则可以对系统资源利用率进行评价。对于如何确定粒度同具体的业务应用以及相关人士所需要关注的指标有很大的关系。

KPIs: 在聚合后为了全面反映系统的运行状况, 根据系统评价需求而加入系统以外的其他数据后形成的复合指标。

服务 KQIs: 不同的 KPIs 聚合会得到不同的服务 KQIs, 是对底层指标的综合和抽象, 比如彩铃业务可利用度。

产品 KQIs: 不同的服务 KQIs 经过聚合后反映整个产品的运行质量, 比如彩铃业务总体运行质量, 这个指标由彩铃业务可利用度、彩铃业务运行效率、彩铃业务准确度、彩铃业务用户满意度等 KQI 组成。

为了解决 TMF 提出的解决方案中存在的问题, 本文改进了关键指标层次(图 2), 可以聚集不同类型、不同意义的指标进行评价, 比如延时和带宽。定义特定业务或者网络 KPI 和 KQI 请参阅文献[2]。

4 K2K 算法

4.1 算法内容

K2K 算法在计算各层指标的相对权重时, 利用了层次分析法^[3]。在 KPI 向 KQI 映射时, 主要采用了 K2K 算法, 因为具体的 KPI 在对 KQI 中贡献有大有小, 根据 AHP 算法可以较精确计算出图 3 中各层兄弟节点的相对权重, 以便求得系统对应健康状况的因子。

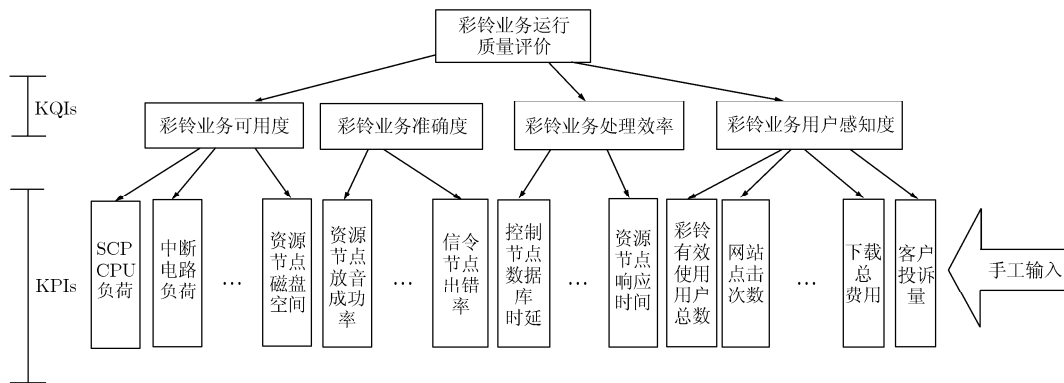


图 3 系统例图

因子定义如下:代表某个指标在一段时间内所展示健康状况的量化,在同层兄弟节点中因子集合经过计算后是固定不变的,除非在层次树中有新的节点加入需要重新计算,对于如何计算因子请参考以下的因子计算部分。结合层次分析法下面着重讨论如何由 KPI 映射到 KQI,以彩铃业务运行质量评价为例。

这里假设一个 KQI 下有 n 个 KPI, 则其相对于上层 KQI 权重为 $(w_1, w_2 \dots w_n)$ 。指标的健康状态一般分为 3 档, 这里分为几档可以根据系统的 SLA^[4](Service Level Agreement)或者根据服务质量等级决定。下面将图 3 中各层指标的质量等级都定义为 3 档: (1)第 1 档代表指标正常。(2)第 2 档代表指标处于亚健康。(3)第 3 档代表某个指标问题比较严重。

对第一档范围指标正常我们可以设初始因子为 1, 然后对于亚健康状态, 根据其权重 $(w_1, w_2 \dots w_n)$ 来计算其因子, 计算过程如下:

指标正常范围为

$$\text{normal} = \sum_{k=1}^n w_k \quad (1)$$

所以其父 KQI 的指标正常状态为 $(0 \sim \text{normal})$ 。

对于某个指标正常由假设可以知道 $(0 \sim \text{normal})$ 为 $(0 \sim 1)$ 。对于指标处于亚健康的计算过程如下所示:

当其中的一个 KPI(比如 CPU 负荷)处在亚健康运行状态的时候, 设其因子系数为 x , 则

$$y = \text{Item}(\text{MIN}(w_1 \dots w_n)) \quad (2)$$

y 代表最小权值的位置。

为求得最小左限引入式(3)

$$\begin{aligned} \text{left_abnormal} &= x \cdot \text{MIN}(w_1 \dots w_n) + \sum_{k=1 \& k! = y}^n w_k \\ &= \text{normal} + 1 \end{aligned} \quad (3)$$

根据式(3)可以推导出 x 值:

$$x = \left| \frac{\text{normal} + 1 - \sum_{k=1 \& k! = y}^n w_k}{\text{MIN}(w_1 \dots w_n)} \right| \quad (4)$$

设 $\text{cf} = x$, 则

$$\text{left_abnormal} = \text{cf} * \text{MIN}(w_1 \dots w_n) + \sum_{k=1 \& k! = y}^n w_k \quad (5)$$

从式(3)可以看出 left_abnormal 一定大于 normal , 做到了对于其父节点不同指标质量范围的合理划分, 并且随着不同 KPI 权重的变化, 其 left_abnormal 能准实时反映所评价指标的真实运行质量状况。

$$\text{right_abnormal} = \text{cf} \sum_{k=1}^n w_k = \text{cf} \quad (6)$$

根据指标健康状况的数量可以得出 N 个因子。可以总结出通用计算因子的公式如下:

$$\text{Factor}(x) = \left| \frac{\text{Previous}_{\text{right}} + 1 - \sum_{k=1 \& k! = y}^n w_k}{\text{MIN}(w_1 \dots w_n)} \right| \quad (7)$$

$\text{Previous}_{\text{right}}$ 为前次计算的右边范围值。 w_k : k 是除了 y 以外的正整数。

K2K 算法目前属于初级研究阶段, 还没有考虑如果 KPI 的采集时间不一样、KPI 的范围划分需要用户的经验值等问题。现在某些组织已经定义了一些具体业务的感知范围的划分度量, 如文献[5]中所述。

4.2 时间复杂度

设评判树有叶子节点数 K_0 , 其上一层为 K_1 , 依次知道上层为 $K_2, K_3 \dots$ 个节点, 设树中总的节点数为 n , 则计算 n 个节点的因子系数时间为 $n - K_0$, 当计算完毕因子后因子系数集保持不变, 除非在层次树中有新的指标加入, 当收集的测量数据向上迭代的时候判别健康度时间复杂度为 n , 则总的所用时间为 $2n - K_0$, 由以上可以明确地得出本算法的时间复杂度为 $O(n)$, 所以对参加运算的指标数量不会进行限制, 能适应较大规模的网络。

4.3 K2K 算法与 TMF 提出聚合算法的对比

下面以智能网方式^[6]彩铃业务可利用度指标为例如表 1 所示。

表 1 彩铃业务可利用度指标集

KQI	KPI
彩铃业务可利用度	SCP(service control point) CPU 负荷
	SCP 硬盘剩余空间率
	SCP 内存剩余率
	SCP IO 吞吐量
	SCP 不可用时长
	SCP 试呼次数
	控制节点资源占用话务量

根据表 1, 利用 4.1 节提到的 K2K 算法可以直接聚合得到业务可利用度, 如表 2 所示。显然由于不能聚合不同量纲不同意义的指标, KQI 的指标数大大增加, 不利于直观判别系统的运行状况。根据上面对比表 1 和表 2 可以看到, K2K 算法对 TMF 所提算法有了很大的改进和提高, 将系统和业务运行状况上升到用户感知层面^[7]。

表2 TMF 算法聚合指标集

KQI	KPI
系统资源利用度	SCP CPU 负荷
	SCP 硬盘剩余空间率
	SCP 内存剩余率
系统负荷	SCP IO 吞吐量
	控制节点资源占用话务量
系统不可利用时长	SCP 不可用时长
外部系统状况	SCP 试呼次数

4.4 K2K 算法性能分析

本文分别以指标数 200, 400, 600, 800, 1000 对算法的性能进行了验证, 并且同 TMF 提出的算法进行了对比, 结果如图 4 所示。

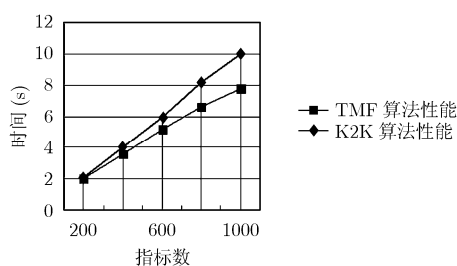


图4 K2K 算法和 TMF 提出算法性能对比图

从图 4 中可以看出, 单纯从性能来看 TMF 提出的算法由于简单性能比较好, 算法所用时间比较少。K2K 算法由于要计算因子系数以及得分计算从而增加了运算时间, 但是由于无论是计算因子系数还是得分计算都是线性的, 对算法的效率影响比较少。从图 4 可以看出, 计算时间随指标数的增长基本是呈线性的。另外, 随着计算机硬件的迅速发展, 硬件处理能力越来越强, 可以得出结论, 本文提出的 K2K 算法能适合大规模网络的需求。

5 结束语

从 K2K 算法的推导过程也可以看出其数学建模同具体的底层指标内容是相分离的, 所以 K2K 算法可以应用于不同行业的指标评价中, 具有广泛的实用性。但是, 在 K2K 算法中没有考虑指标之间的相关性分析^[8], 导致在计算得分的时候出现分数偏高或者偏低。在后续研究中希望能考虑以下几个问题: (1)考虑指标之间的相关度。(2)对新指标具有

自学习的功能。(3)对指标具有通过学习自适应分类^[9]的功能。

上面的几点涉及到了人工智能以及数据挖掘^[10]方面的一些内容, 其对指标的获取以及指标的评价的准确度具有深刻的意义。

参考文献

- [1] Telemanagement Forum, SLA management handbook concepts and principles. Vol 2, <http://www.tmforum.org/>, 2005.01: 60-70.
- [2] Telemanagement Forum, GB923 wireless service measurements handbook. <http://www.tmforum.org/>, 2004. 03: 42-60.
- [3] 许树柏. 层次分析法原理[M]. 天津:天津大学出版社, 1988: 60-90.
- [4] Park Long-Tae, Baek Jong-Wook, and Hong Woon-Ki. Management of service level agreements for multimedia Internet service using a utility model[J]. *IEEE Communications Magazine*, 2001, 39(5): 100-106.
- [5] ITU-T G.1010-2001, End-user multimedia QoS categories[S]. Nov.2001: 8-10.
- [6] 中国移动研究院. 彩铃业务总体技术要求“v200 下版”[S]. 中国移动通信公司彩铃业务技术要求-2004, 2004. 09.
- [7] Stefic R and Prib N. Measurement and analysis of users' perception of QoS for IP telephony service[C]. 7th International Conference on Telecommunications-ConTel, Zagreb Croatia, June 11, 2003: 505-512.
- [8] Qu G, Hariri S, and Yousif M. A new dependency and correlation analysis for features[J]. *IEEE Trans. on Knowledge and Data Engineering*, 2005, 17(9): 1199-1207.
- [9] Freitas A A. A survey of evolutionary algorithms for data mining and knowledge discovery[C]. ACM Advances in evolutionary computing:Theory and applications, Jan. 2003: 819-845.
- [10] Chen Ming-Syan, Han Jiawei, and Yu PS. Data Mining: An overview from a database perspective[J]. *IEEE Trans. on Knowledge and Data Engineering*, 1996, 8(6): 866-883.

倪萍: 男, 1978年生, 博士生, 研究方向为宽带通信网与网络管理。

廖建新: 男, 1965年生, 教授, 博士生导师, 研究方向为移动智能网、宽带智能网。