

面向呼叫中心的知识导航系统研究

吴江宁*, 王晓欢

(大连理工大学 系统工程研究所, 辽宁 大连 116024)

摘要: 从系统构建的角度分析了当前呼叫中心存在的问题,提出了面向呼叫中心的知识导航系统构建方法,旨在改善现有呼叫中心的整体性能.该系统分为基于主题地图的显性知识导航,及基于人际网络的隐性知识导航两个部分.构建该系统的具体方法包括主题地图构建、人际网络构建、知识点关联、按能力和关系导航业务代表,及对业务代表个人特征的评价.结合中国移动通信公司呼叫中心的实际情况,设计并实现了面向该领域的知识导航原型系统,并对构建该系统的具体方法进行了评估.

关键词: 知识导航;主题地图;人际网络;呼叫中心

中图分类号: N945.12 **文献标志码:** A

0 引言

呼叫中心(call center)是企业建立的基于通信网络和计算机网络的服务中心,它是企业与其客户进行信息交互的渠道^[1].呼叫中心系统大致包含以下几个子系统:自动语音/传真应答子系统(IVR)、计算机电信集成子系统(CTI)、自动呼叫分配子系统(ACD)、人工坐席台子系统、系统管理子系统及数据库子系统,这些子系统相互依赖、相互作用构成了现代呼叫中心系统.目前,大量研究都集中在完善以上子系统的功能层面,例如徐雅斌等^[2]提出了基于语音卡和 CTI 技术的通用呼叫中心系统框架,并改进了传统排队模型,旨在提高客户服务质量;朱仁祥等^[3]改进了传统的 ACD 排队模型,并对系统整体性能和 ACD 子系统进行了统计数据分析;吴佳骥等^[4]提出了一种多技能呼叫中心人力需求的计算模型,提供了新的呼叫路由解决方案;谢燕等^[5]引进数据仓库技术,提高了呼叫中心处理大量历史数据的效力和效率;刘秋梅等^[6]引入数据挖掘技术,构建了基于呼叫中心的数字参考服务模型,旨在提高呼叫中心系统发现潜在用户的能力.然而,目前对呼叫中心系统的研究仍然存在一些不足,主要体现在:

(1)IVR 虽然可以较好地回答客户查询,但

与数据库子系统中存储的海量知识相比,其所解答问题的知识结构比较单一,涉及范围较窄,缺乏知识应答的灵活性;

(2)ACD 虽然可以较好地实现呼叫排队功能,但它忽略了客户需求与业务代表专业知识之间的关联,造成了指派的业务代表的个人能力与实际需求互不匹配,影响客户服务质量;

(3)业务代表在处理客户需求时,常常遇到个人无法解决的问题,有时通过人工坐席台子系统,也无法找到合适的求助对象.

基于以上分析,本文提出通过构建面向呼叫中心的知识导航系统,使其与呼叫中心其他子系统协同工作,进而改善呼叫中心的整体性能,提高客户服务质量.

1 知识导航技术的研究现状

根据知识的分类,知识导航可以分为显性知识导航和隐性知识导航.传统的显性知识导航技术包括基于关键字的检索技术、基于分类目录的导航与检索技术、基于概念的导航与检索技术^[7].这些方法很少考虑信息之间的语义相关性,因此知识导航效率较低.近年来,语义网的发展加快了对显性知识导航技术的研究.目前,基于语义

②提取业务代表解决客户需求的历史记录, 将其与主题地图的知识点相关联, 进而实现通过主题地图来查找能够解决某类实际问题的业务代表信息;

③提取业务代表的基本信息, 结合对业务代表进行问卷调查的结果, 构建人际网络, 提供按业务能力与业务代表之间的关系进行排序的结果.

(2)与 ACD 子系统关联

①当客户选择自动应答服务时, ACD 子系统将匹配得到的知识点所对应的客户需求解答方案传递给 IVR;

②当客户选择人工服务时, ACD 子系统将匹配得到的知识点所对应的能够解决该问题的业务代表列表反馈给人际网络.

(3)与人工坐席台子系统关联

①当人工坐席台子系统发出业务代表查询请求时, 人际网络提供按能力和关系两种类型划分的业务代表列表;

②人际网络将这一结果返回给人工坐席台子系统, 供其进一步决策.

3 知识导航系统中的技术及方法

针对中国移动通信公司呼叫中心的实际情况, 本章给出了知识导航的具体方法. 在此基础上, 设计并实现了面向中国移动通信公司呼叫中心的知识导航原型系统. 该系统主要面向客户服务领域, 可以实现显性知识导航、知识点关联及业务代表导航 3 个主要功能.

由于本文所涉及的客户需求信息存在知识结构复杂和信息量大的特点, 在检索历史客户需求信息数据库进而进行显性知识导航时, 为了提高信息检索的效率, 引进了基于向量空间模型的相似度计算方法, 即利用向量空间模型 (vector space

model, VSM)^[16] 进行文本特征表达, 用 tf-idf (term-frequency inverse-document-frequency) 进行特征词赋权, 用倒排文档进行索引, 用余弦夹角进行距离度量, 用准确率和召回率进行检索评价^[17].

下面结合实例, 具体介绍知识导航的技术及方法.

3.1 主题地图构建

主题地图是一种新兴的知识组织和知识表示技术, 它作为一项国际标准被定义在 ISO13250 中. 它能够描述信息资源的知识结构, 并确定相关知识概念所在资源的位置. 主题地图由主题、关联、事件 3 个基本要素构成^[10]. 针对中国移动通信公司呼叫中心的实际情况, 本文从主题提取、事件回挂及关联分析 3 个方面构造领域主题地图.

(1)主题提取

根据文本切分算法^[18]对历史客户需求信息进行分词; 从分词结果中选取具有代表性的词条, 运用概念聚类算法^[19]对词条进行聚类; 根据聚类结果, 确定主题地图的主题及主题类别. 这里, 定义主题类别为 T_i , T_{ij} 为主题类别 T_i 的下位关系主题, T_{ijk} 为主题类别 T_{ij} 的下位关系主题 (其中 i, j, k 为自然数), 依此类推.

本例从某市移动通信公司呼叫中心的历史客户需求信息库中抽取 500 条具有代表性的信息作为测试文本集, 应用上述的文本切分算法可以得到 5 228 项分词结果; 依据一定原则, 从中选出 210 项词条作为初步特征词; 根据聚类算法, 确定 5 个主题类别及 87 个主题概念.

如图 2 所示的主题地图中, 各个主题概念可被依次定义为 T_1 = “服务项目”, T_{11} = “数据通信业务”, T_{12} = “话费”, T_{111} = “彩铃业务”, T_{112} = “GPRS 业务”, T_{121} = “欠费”, T_{122} = “信息费”.

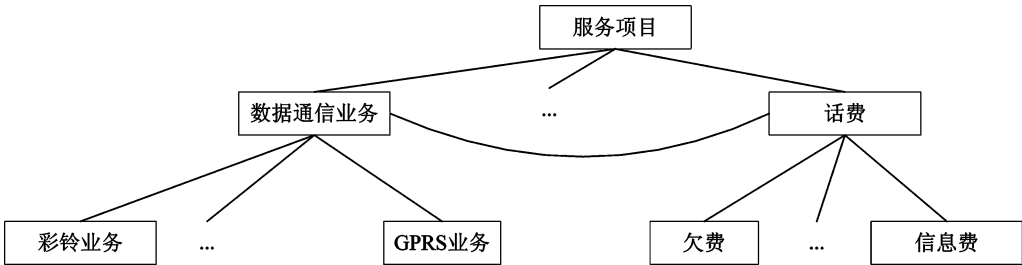


图 2 部分主题地图知识层

Fig. 2 Knowledge level of partial topic map

(2)事件回挂

将与主题相关的历史客户需求信息、解决方案及能够解决该问题的业务代表信息回挂到相应的主题上. 其具体步骤如下.

第一步:主题地图向量化. 处于不同层次的主题包含的信息量不同,即层次越高,信息越抽象;层次越低,信息越具体. 因此,不同的主题应该被赋予不同的权值,不妨设 w 是主题的权值,则计算公式为

$$w = h / (H \cdot K^r)$$
 (1)

其中 h 是当前主题在主题地图中的高度; H 是主题所在分支的层次高度(根主题为第 0 层); r 为当前主题与其所在分支的最底层主题之间的距离; K 为常量,通常取 $K = 2$. 这样,主题地图可以表示成一个一维向量,即 $T = (w_1 \quad w_{11} \quad \cdots \quad w_{1i} \quad \cdots \quad w_{1jk} \quad \cdots \quad w_i \quad \cdots \quad w_{ij} \quad \cdots \quad w_{ijk} \quad \cdots)$.

以图 2 中的主题地图为例,各个主题概念的权值如表 1 所示.

表 1 “服务项目类”主题概念树权重一览表

Tab. 1 Table of "service" topic concept tree's weights

主题概念	层次	权重
服务项目类	0	0
数据通信业务,话费,...	1	0.25
彩铃业务,GPRS 业务,欠费,信息费,...	2	1

则该主题概念树可以表示成一维向量,即 $T = (0 \quad 0.25 \quad 1 \quad 1 \quad 0.25 \quad 1 \quad 1)$.

第二步:构建主题 - 文本索引矩阵. 记历史客户需求信息文本为 D_m ,与主题地图中的主题 T_{ijk} 构成主题 - 文本矩阵,矩阵中的元素记为 t_{mijk} ,是主题 T_{ijk} 在文本 D_m 中出现与否的判据,其取值可

定义如下:

- ① $\forall m \in M$,若 $T_i \in D_m$,则 $t_{mi} = 1$ 或 $T_i \in D_m$ 且 $T_{ij} \in D_m$,则 $t_{mi} = 0, t_{mij} = 1$;
- ② $\forall m \in M$,若 $T_{ij} \in D_m$,则 $t_{mij} = 1$ 或 $T_{ij} \in D_m$ 且 $T_{ijk} \in D_m$,则 $t_{mij} = 0, t_{mijk} = 1$.

其余依此类推.

第三步:事件回挂. 首先,根据主题-文本索引矩阵,将历史客户需求信息回挂到主题地图的各个主题上;然后,从对应历史客户需求信息的解决方案中提取解决该问题的业务代表信息,将其作为主题地图的事件信息直接回挂到相应的问题(主题)上,如图 3 所示.

(3)关联分析

根据企业/组织的实际情况,分析上述主题之间及主题与主题类别之间的语义关联,并将这些关联标识在主题地图中. 本例共涉及两种语义关联,即包含关系和影响关系,如图 4 所示.

以上是主题地图的构建方法. 在实现原型系统时,根据上述方法,利用本体编辑工具 protégé 绘制了相应的领域主题地图,如图 5 所示. 这样,业务代表可以通过导航机制浏览此主题地图,实现领域显性知识的导航.

3.2 人际网络构建

业务代表是呼叫中心系统中不可或缺的重要组成部分,他们在人工客户服务中发挥着重要的作用. 随着客户服务业务的开展,客户需求的复杂性与日俱增,当业务代表遇到自己无法解决的客户需求时,他会尝试求助其他的业务代表. 这种行为往往基于一定的人际关系原则,如徒弟求助师傅等. 本文引入人际网络^[20]来描述业务代表之间

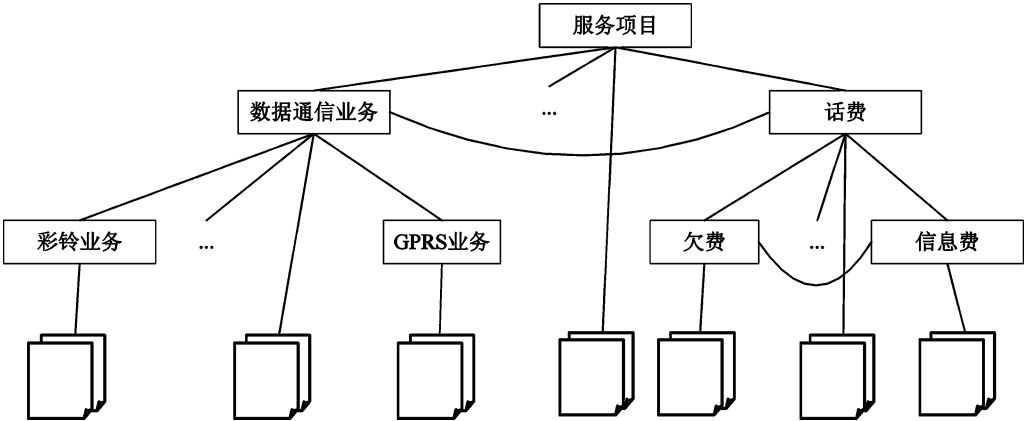


图 3 部分主题地图的事件回挂

Fig. 3 Occurrence appending of partial topic map

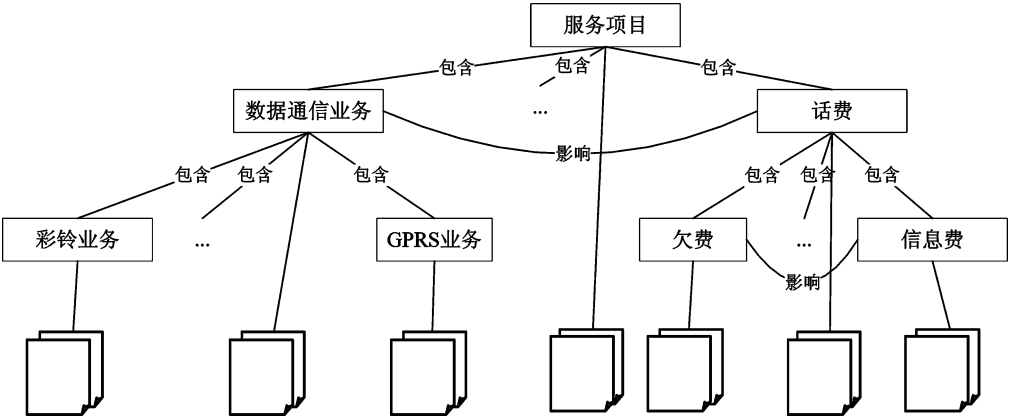


图 4 “服务项目类”部分主题地图
Fig. 4 Partial topic map of "service"

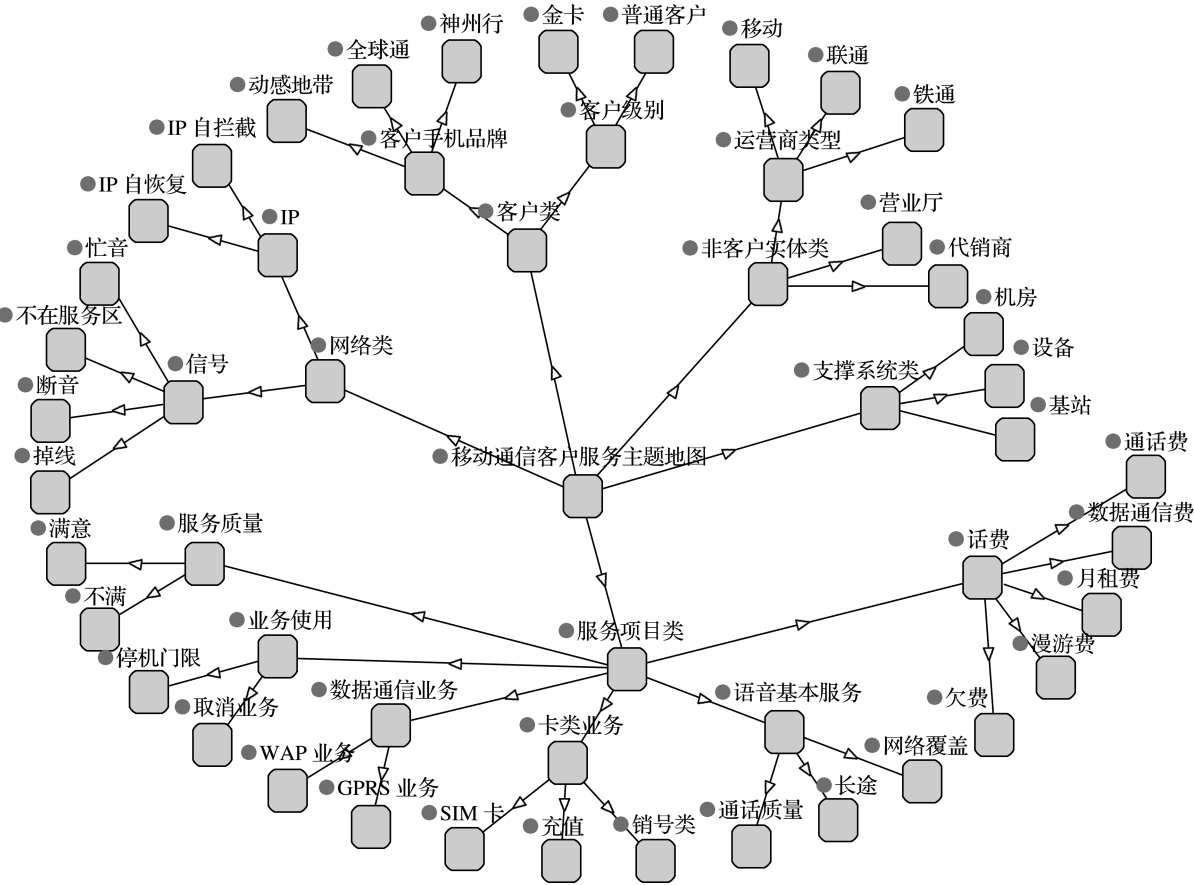


图 5 中国移动通信公司客户服务领域的主题地图
Fig. 5 Domain topic map for customers' services of China Mobile Communication Corporation

的关系. 由于人是具有社会性的个体, 人与人之间的关系往往错综复杂. 为了描述业务代表之间的复杂关系, 并依此选择合适的业务代表作为求助对象, 本文提出了具有多关系的人际网络映射算法, 旨在实现有效的隐性知识导航. 其具体步骤如下:

- (1) 将业务代表之间的多种关系 R_i 绘制成多个具有单种关系的人际网络 N_i .
- (2) 根据业务代表之间关系的重要程度, 赋予 N_i 相应的权值 λ_i .
- (3) 将不同的人际网络 N_i 映射到一个新的人际网络上. 其映射原则为各网络中的点和点相

互重合;边的权值取所有重合边中权值最大的一项.

新网络即为具有多关系的人际网络.

问卷调查是一种最普遍应用的人际关系获取方法,因此本文采用问卷调查的方式来提取业务代表之间的人际关系,并根据上述算法确定人际

网络.

以 10 位业务代表为例,假设他们之间具有两种人际关系.图 6 中(a)、(b)分别代表权值为 λ_1 和 λ_2 的业务代表关系网络,不妨设 $\lambda_1 = 2\lambda_2$. 依据上述算法,可以得到一个具有两种关系的新业务代表网络,如图 6(c)所示.

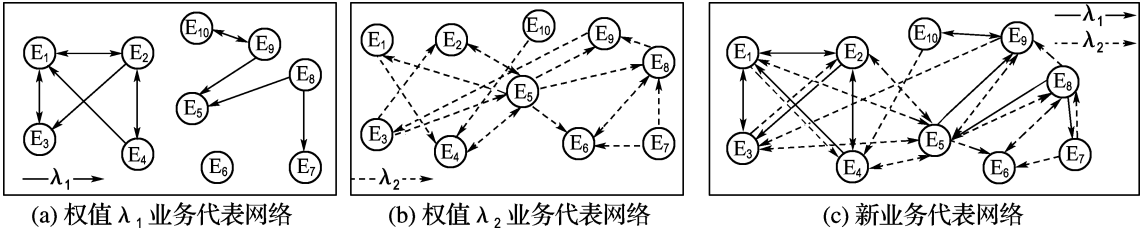


图 6 业务代表关系映射示意图

Fig. 6 Experts' relationships mapping process

3.3 知识点关联

本文利用向量空间模型计算客户需求信息与该服务领域知识结构之间的相似程度,求解相应的关联知识点.其具体步骤如下:

(1) 根据 3.1,将历史客户需求信息及新客户需求信息表示成向量的形式,即

$$D_m = (\begin{matrix} w_{m1} & w_{m11} & \cdots & w_{m1i} & \cdots & w_{m1jk} & \cdots \\ & w_{mi} & \cdots & w_{mij} & \cdots & w_{mijk} & \cdots \end{matrix}) \quad (2)$$

$$Q = (\begin{matrix} w_{q1} & w_{q11} & \cdots & w_{q1i} & \cdots & w_{q1jk} & \cdots \\ & w_{qi} & \cdots & w_{qij} & \cdots & w_{qijk} & \cdots \end{matrix}) \quad (3)$$

(2) 利用向量夹角余弦公式计算新客户需求信息与全部历史客户需求信息的相似度,即

$$\begin{aligned} sim(D_m, Q) &= \cos \theta = \\ &= (\begin{matrix} w_{s1}w_{q1} + \cdots + w_{si}w_{qi} + \cdots + \\ w_{sj}w_{qj} + \cdots + w_{sjk}w_{qijk} \end{matrix}) / \\ &= (\sqrt{w_{s1}^2 + \cdots + w_{sjk}^2} \times \\ &\quad \sqrt{w_{q1}^2 + \cdots + w_{qijk}^2}) \end{aligned} \quad (4)$$

(3) 设定阈值 γ ,当 $sim(D_m, Q) \geq \gamma$ 时,将得到的历史客户需求信息对应的主题地图中的主题(知识点)提取出来,这就是知识点关联的结果.

本文将 10 条新客户需求信息与主题地图中存储的历史客户需求信息进行匹配,从而将新客户需求信息与主题地图知识点进行关联,通过信息检索的评价指标,即准确率和召回率,对原型系统的性能进行评价.设定阈值 $\gamma = 0.6$.

如表 2 所示,系统中平均 84.64% 的知识点关联比较准确,而平均 69.68% 的知识点关联比较全面.实验中,第 3 条和第 9 条的知识点关联结

果没有满足阈值限定的要求,视为知识点关联失败.则系统知识点关联成功率为 80%,成功关联后,系统中平均 89.27% 的知识点关联比较准确,而平均 74.23% 的知识点关联比较全面.这说明基于主题地图的信息检索原型系统的性能良好.

表 2 知识点关联实验结果

Tab. 2 Testing results of knowledge associations

序号	检出文本	检出正确文本	实际正确文本	准确率/%	召回率/%
1	47	44	71	93.62	61.97
2	47	47	65	100	72.31
3	5	5	14	100	28.57
4	5	5	6	100	83.00
5	10	7	8	70.00	87.50
6	9	7	9	77.78	77.78
7	22	16	23	72.73	69.57
8	2	2	3	100	66.67
9	62	20	27	32.26	74.07
10	3	3	4	100	75.00
系统平均值				84.64	69.68
(成功关联后系统平均值)				(89.27)	(74.23)

3.4 按能力查找业务代表

本文将业务代表的培训次数 TT (training times) 及工作效率 WE (work efficiency) 作为考核业务代表能力 DA (deputy ability) 的指标. $DA(D_i, t)$ 表示在 t 时刻,业务代表 D_i 的能力,不妨设业务代表初始工作的时刻为 t_0 ,则可得如下公式:

$$DA(D_i, t) = f[a_1 TT(D_i, t) + a_2 WE(D_i, t)] \quad (5)$$

其中 $a_1 + a_2 = 1$.

$TT(D_i, t)$ 通过业务代表 D_i 在时间 $(t - t_0)$ 内参加过的相关培训的次数 $x(D_i, t)$ 来衡量,即

$$TT(D_i, t) = \frac{x(D_i, t)}{t - t_0} \quad (6)$$

$WE(D_i, t)$ 通过业务代表 D_i 在时间 $(t - t_0)$ 内解决问题的次数 $y(D_i, t)$ 来衡量,即

$$WE(D_i, t) = \frac{y(D_i, t)}{t - t_0} \quad (7)$$

这样就可得到在不同时刻每位业务代表的能力.人工坐席台子系统可以根据业务代表的能力排序,指定合适的业务代表来解决客户提出的新需求.

3.5 按关系查找业务代表

当业务代表遇到自身无法解决的问题时,往往会求助与他有某种关系的其他业务代表.利用 3.2 中的具有多关系的人际网络,可以实现按关系对业务代表的查找.具体过程如下:

(1) 根据实际查找需要,调整具有多关系的人际网络中各边的权值,即关系越重要,权值越大;反之亦然.

(2) 利用最短路径原理经典算法 Dijkstra 标号法,计算一个业务代表到其他业务代表之间的距离,并得到业务代表求助路径.

根据 3.4 及 3.5 介绍的方法,人工坐席台子系统可以依据业务代表的能力及其相互之间的关系,选择合适的业务代表来帮助当前业务代表解决客户需求,从而提高客户服务质量.

3.6 业务代表个人特征评价

利用社会人际学软件 SociometryPro 可以对具有多关系的人际网络中的各位业务代表的个人特征进行评价.管理者根据评价结果,因人而异,制定相应的提高业务代表工作质量的方法.图 7 为对 10 位业务代表进行个人特征评价的结果,评价指标包括业务代表在团队中的重要性、影响力及满意度.关于 SociometryPro 工具的介绍及指标计算公式见文献[21].

由图 7 可知:(1)业务代表 E_7 、 E_{10} 的重要性最低.说明团队中其他成员对这两位业务代表的求助最少,管理者可以借此考核这两位业务代表的工作能力及人际关系,找出他们不被求助的具体原因,并加以解决.

(2)业务代表 E_5 的影响力最高.说明该业务代表是最有可能影响其他人的人.因此,管理者可以根据该业务代表的这一特点,赋予他相关的工

作,从正面引导该业务代表为团队服务.

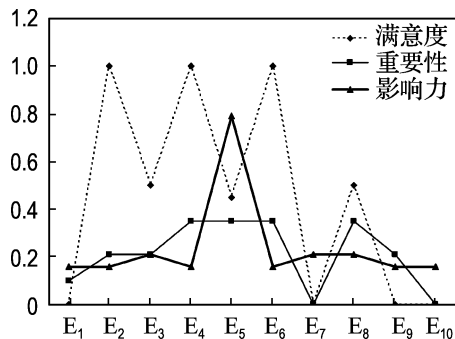


图 7 对 10 位业务代表个人特征评价结果

Fig. 7 Evaluation results of 10 experts' performance

(3) 业务代表 E_2 、 E_4 、 E_6 的满意度最高,而 E_1 、 E_7 、 E_9 、 E_{10} 的满意度最低.说明前 3 位业务代表对工作的积极性较高,而后 4 位业务代表对工作的积极性较低.因此,管理者应该激励前 3 位业务代表,使其进一步发挥他们的积极作用,多与后 4 位业务代表进行沟通,找到他们工作积极性较低的原因,并加以解决.

4 结 语

本文设计并实现了面向中国移动通信公司呼叫中心的知识导航系统,并对该系统的性能进行了评估.结果表明该系统具有较好显性知识导航、知识点关联及业务代表导航的能力,在一定程度上可以优化 IVR 的自动应答效果,改善人工坐席台子系统的工作效果;另外,利用 SociometryPro 工具对业务代表个人特征进行评价,可为管理者制定措施提供依据,进一步优化人工坐席台子系统的工作效果,最终提高呼叫中心系统的整体性能.

但是,目前对隐性知识导航的研究只局限于业务代表的能力及业务代表之间的关系层面上,而对业务代表的个人知识涉及较少,且本文对人际网络的划分是基于问卷调查的方式确定的,该方法的误差性较大.因此,今后的研究将侧重管理业务代表的个人知识,并开发更有效的人际网络划分机制,借以提高知识导航系统的性能,从而最终提高客户服务的质量.

参考文献:

- [1] 张志远,张忠能,凌君逸.基于呼叫中心的客户管理系统在制造业的应用[J].计算机工程,2004,30(增

- 刊):427-429
- [2] 徐雅斌,张云帆. 基于 CTI 技术的呼叫中心的设计与实现[J]. 计算机工程, 2007, **33**(5):232-236
- [3] 朱仁祥,朱翼隼,方基奎. 重试, 反馈 M/M/s/k 排队的呼叫中心性能分析[J]. 系统工程学报, 2006, **21**(6):613-620
- [4] 吴佳骥,黄刘生,吴俊敏,等. 一种多技能呼叫中心人力需求的计算模型[J]. 上海交通大学学报, 2007, **41**(7):1169-1175
- [5] 谢 燕,郑有才,张立勇,等. 数据仓库技术在呼叫中心中的应用[J]. 计算机工程与设计, 2006, **27**(21):4150-4160
- [6] 刘秋梅,郑耿忠. 呼叫中心数据挖掘在数字参考服务中的应用研究[J]. 情报杂志, 2006, **25**(5):68-72
- [7] BEIER J, TESCHE T. Navigation and interaction in medical knowledge spaces using topic maps [J]. **International Congress Series**, 2001, **1230**:384-388
- [8] ALANI H, HOSER B, SCHMITZ C, *et al.* Semantic network analysis [C] // **Proceedings of the 2nd Workshop on Semantic Network Analysis**. Montenegro:[s n], 2006
- [9] MILLER E. An introduction to the resource description framework [J/OL]. D-Lib Magazine, <http://www.dlib.org/dlib/may98/miller/05miller.html>
- [10] 李善平,尹奇桦,胡玉杰,等. 本体论研究综述[J]. 计算机研究与发展, 2004, **41**(7):1041-1052
- [11] PEPPER S. The TAO of topic maps. [2006-06-20]. <http://www.ontopia.net/topicmaps/materials/tao.html>
- [12] 邓三鸿,金 莹,杨建林. 学科知识地图的构建[J]. 情报学报, 2006, **25**(1):3-8
- [13] 郑邦坤. 隐性知识信息组织研究[J]. 情报杂志, 2004, **23**(7):63-68
- [14] 殷国鹏,莫云生,陈 禹. 利用社会网络分析促进隐性知识管理[J]. 清华大学学报(自然科学版), 2006, **46**(s1):964-969
- [15] 王众托. 系统工程引论[M]. 北京:电子工业出版社, 2006
- [16] SALTON G, WONG A, YANG C S. A vector space model for automatic indexing [J]. **Communications of the ACM**, 1975, **18**(11):613-620
- [17] 杨小平,丁 浩,黄都培. 基于向量空间模型的中文信息检索技术研究[J]. 计算机工程与应用, 2003, **39**(15):109-111
- [18] 姜韶华,党延忠. 基于长度递减和串词统计的文本切分算法[J]. 情报学报, 2006, **25**(1):74-79
- [19] WU Jiang-ning, TIAN Hai-yang, YANG Guang-fei. A multilayer topic-map-based model used for document resources organization [M] // **Lecture Notes in Control and Information Sciences:Vol 344. Intelligent Control and Automation**. Berlin:Springer, 2006
- [20] 苏俊燕,孔令江,刘慕仁. 加权网络上的舆论演化模型研究[J]. 广西师范大学学报(自然科学版), 2006, **24**(2):1-4
- [21] LeDis Group. SociometryPro 2.3 [Z/OL]. <http://www.ledisgroup.com>

Research on call-center-oriented knowledge navigation system

WU Jiang-ning*, WANG Xiao-huan

(Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: From the perspective of system construction, a novel knowledge navigation system was developed for solving the problems of the current call centers so as to improve their performance. Such system consists of two sub-systems: one is the topic map based explicit knowledge navigation system and the other is the interpersonal network based tacit knowledge navigation system. An approach to implementation of the system is specified in five aspects: topic map construction, interpersonal network construction, knowledge node linking, ability and relationship oriented representative expert navigation and representative expert's performance evaluation. A case for the call center of China Mobile Communication Corporation was deeply investigated. Accordingly a prototype system was designed and developed for the domain knowledge navigation. Meanwhile, some experiments were carried out to prove the effectiveness of approaches to the system construction.

Key words: knowledge navigation; topic map; interpersonal network; call center