基于文本事件网络自动摘要的抽取方法

杨竣辉[1,2] 刘宗田[1] 刘 炜[1] 苏小英[1]
(上海大学计算机工程与科学学院 上海 200072)[1] (江西理工大学信息工程学院 赣州 341000)[2]

摘 要 将文本按事件方式进行表示,把事件作为基本语义单元来构建事件本体。根据事件间的关系构建事件网络有向图能较好地表达文本的语义信息及事件间的关系重要程度。利用PAGERANK算法测算事件网络图中各节点对应事件的重要度并进行排序,按事件发生的时间顺序,输出事件对应的原语句作为摘要。实验结果表明,基于事件网络的文本自动文摘方法抽取出的摘要效果较好。

关键词 文本表示,事件本体,事件网络,PAGERANK
中图法分类号 TP311 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.3.043

Extraction Method of Text Summarization Based on Event Network

YANG Jun-hui[1,2] LIU Zong-tian[1] LIU Wei[1] SU Xiao-ying[1]
(School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China)[1]
(School of Information and Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China)[2]

Abstract Text was expressed by the means of event, and event ontology was built by using event as the basic semantic unit. According to the relationship between events, we built event network direct diagram which can express more semantic information of the text and describe the importance of relationship between events. The importance degree of event of the event network corresponding to each node was calculated and ranked by using the PAGERANK algorithm. According to the time sequence of events, event corresponding primitives were exported as abstract. The experimental results show that automatic summary based on the event network method has better performance.

Keywords Text representation, Event ontology, Event-Network, PAGERANK

1 引言

自动摘要技术从20世纪50年代开始兴起,最初是以统计学为支撑,依靠文章中的词频、位置等信息为文章生成摘要,主要适用于格式较为规范的技术文档[1]。自动摘要技术的本质是信息的挖掘和信息的浓缩,是实现自然语言理解的重要标志之一,对自动摘要技术的研究有助于理解自然语言文本,获取知识模型。

近年来,随着自动文摘研究的深入,研究者们不断提出各种方法。其中,Lin等人[2]在1999年提出另一种摘要方法。在这种方法中,他们假设文章中用于摘要抽取的各种特征是相互关联的,并使用了决策树(decision trees)而不是贝叶斯分类模型对句子打分,将抽取得分最高的部分句子作为文章摘要。

Chang-jin Jiang等[3]通过识别组合词和段落聚类实现中文自动摘要。首先根据词或短语的频率、词性、位置和长度计算它们的权重,在此基础上计算句子的权值;然后将相邻的段落依据相似度聚到相同类或不同类中;最后根据类中句子的权值选择摘要句组成摘要。

Pei-ying等[4]提出了一种基于句子聚类和抽取的自动摘要方法。首先对文本中的句子依据语义距离进行聚类;然后用基于多特征融合的方法计算类中每个句子的权重;最后通过一定规则抽取句子组成摘要。

Munesh Chandra等[5]提出一种基于统计学方法的自动文摘,在统计学的基础上采用kmixture概率模型建立特征词的权重值、识别词之间的语义关系,对提取特征词对应的句子根据权重排名确定摘要。

Erkan[6]提出了基于图进行文本处理的方法,将文本划分为句子集合,构造以句子为顶点的图,利用图计算句子的显著度(salience),并根据显著度抽取句子。

以上学者所提出的方法在实验过程中均具有一定的效果,但方法以获取词或语句的特征为主,这种方法虽然容易实现,但缺乏对词与词之间语义关联的考虑,使得对篇幅较长的文档所生成的摘要难以覆盖文章的主要内容,容易形成大量的冗余。另一种方法是基于图模型的方法,这种方法之所以有效,是因为它通过迭代的计算能有效地获取图的全局信息,也即文本的全局信息,从而在判断句子重要程度上更为准确。相比采用一系列特征为句子打分的方法,这种方法有更好的

到稿日期:2014-04-27 返修日期:2014-06-23 本文受国家自然科学基金项目(61273328,61305053)资助。
杨竣辉(1981-),男,博士生,副教授,主要研究方向为知识表示、自然语言处理和Web数据挖掘等;刘宗田(1946-),男,教授,博士生导师,主要研究方向为人工智能和软件工程等;刘 炜(1978-),男,博士,副研究员,主要研究方向为语义本体、知识表示等;苏小英(1980-),女,博士生,讲师,主要研究方向为数据挖掘、自然语言处理等。

鲁棒性，同时它通常不需要预先设定任何参数。但 Erkan 所提出的基于图模型的方法只考虑了句子与句子之间的上下文关系，并没有考虑到句子与句子间的相似度，这将低估相似度高的语句的重要程度。

美国佛罗里达州大学的 Zwaan[7]将每个单句等同为一个“事件”。在 ACE 评测会议中，“事件”[8]被描述为一个动作的发生或状态的变化。一些学者开始认识到以事件作为知识的基本单元更能反映客观世界知识的运动性实质，并将其用于文本的自动摘要及事件本体的研究。

刘茂福等[9]提出一种基于事件项语义图聚类的多文档摘要方法，即将文档中的动词和动名词看成事件项，然后对事件项进行聚类，最后通过抽取包含代表事件项的句子生成摘要。

韩永峰等[10]提出基于事件抽取的网络新闻多文档自动摘要的方法，通过二元分类器辨析出文本中的事件和非事件；然后通过聚类将文档原来以段落或句子为单位的物理划分转化为以事件为单位的内容逻辑划分，最后通过主旨事件抽取、排序及润色生成摘要。

SS Ge[11]提出基于 SNMF 的句子聚类算法并通过加权图模型考虑句子之间的关系对语句进行排序，最终形成摘要。

Peter Thwaites[12]提出构建事件链图(CEG)作为一种辅助的图形化模型的分析表示事件的因果贝叶斯网络问题。

Zhong[13]提出基于事件关系图的文本事件重要度排序方法，其通过构造事件影响因子矩阵来描述文本集合中事件之间的关联强弱，阐述事件影响关系识别文本集合中重要事件的方法。

综上所述，尽管已有学者提出以事件为单元勾画文档的内容，甚至通过外部语义资源获取的事件项语义关系，并通过一定的方法实现文档的自动摘要，但这些方法多数是以事件与事件间相关程度作为判断句子重要程度的主要因素。有的文献虽然提到以段落或句子为单位事件，计算单位事件到中心事件的距离的长短来确定抽取主旨事件，但还是没有考虑事件间的关联在事件的发生过程中的作用。

本文在常用方法研究文本摘要的基础上，结合对事件的定义和形式化表示，提出了将文本按事件方式进行表示，把事件作为基本语义单元，并根据事件间的关系构建有向的事件网络，采用有向图模型的方法反映文本事件之间的结构关系，较好地表达文本的语义信息，弥补词频特征刻画的不足之处。利用 PAGERANK 测算各节点对应的事件的度来表现事件间的关系重要程度并进行排序后，按事件发生的时间顺序，输出事件对应的原语句作为摘要。最后，通过实验证明了以事件间存在的关系为基础构建事件网络的文本自动文摘方法所抽取出的摘要质量更高。

2 事件及事件网络相关概念定义

定义 1 事件(Event)，指在特定时间和环境下发生的，由若干角色参与的，表现出一些动作特征的一件事情[14]。

事件由六元组 $E=(O,A,T,V,P,L)$ 表示，定义为一个六元组结构。其中，事件六元组中的元素称为事件要素，分别表示对象、动作、时间、环境、断言、语言表现。

事件类(Event Class)，指具有共同特征的事件的集合，用 EC 表示：

$$EC=(E,C_1,C_2,\cdots,C_6) \tag{1}$$

其中，E 是事件的集合，称为事件类的外延；O,A,T,V,P,L 称为事件类的内涵，分别是 E 中的每个事件在对应要素上具有的共同特性的集合。

定义 2(事件本体) 事件本体是共享的、客观存在的事件类系统模型的明确的形式化规范说明，表示为 EO。事件本体的逻辑结构可定义为一个三元结构 $EO:=\{ECS,R,Rules\}$，其中 ECS (事件类集合)是所有事件类的集合；R(事件类之间的关系集合)包括事件类之间的分类关系和非分类关系，分类关系可构成事件类的层次，非分类关系上标明关系种类名，如图 1 所示。

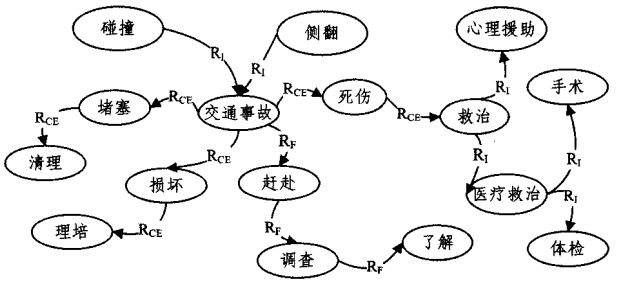

图 1 突发事件本体模型(部分)

定义 3(事件间的非分类关系)

(1)组成关系。如果一个整体 e 可以分解为若干子件 e_i ($i\in[1,n]$)，则称它们之间具有组成关系，称事件 e_i 是事件 e 的组成部分，表示为 $RI(e_1,e_2)$。

(2)因果关系。事件 e_1 的发生导致了事件 e_2 的发生，发生的概率大于给定的阈值，则称两事件间具有因果关系，表示为 $RCE(e_1,e_2)$。

(3) 跟随关系。在一定时间段内，事件 e_2 可能跟随事件 e_1 之后而发生，跟随发生的概率大于给定的阈值，则称两事件间具有跟随关系，表示为 $RF(e_1,e_2)$。

(4)并发关系。在一定长度的时间段内，事件 e_1 和事件 e_2 同时或先后发生，同时或先后发生的概率大于给定的阈值，则称两事件间具有并发关系，表示为 $RC(e_1,e_2)$。

定义 4(事件相似度) 指事件的相似程度，通常用区间[0,1]之间的值来表示。

设事件集合存在任意两个事件 e_i 和 e_j，根据事件要素对应的相似度计算事件的相似度，则有：

$$SIM(e_i,e_j)=\sum_{k=1}^{6}w_k s(e_{ik},e_{jk}),k=(o,a,t,v,p,l) \tag{2}$$

其中，$SIM(e_i,e_j)$是指 e_i,e_j 之间的相似度，e_{ik} 表示事件 e_i 的第 k 个要素，e_{jk} 表示事件 e_j 的第 k 个要素，w_k 表示事件各要素在计算事件相似度时的权重，$\sum w_k=1$，事件相似度的权重介于[0,1]。显然 e_i,e_j 之间的相似度的区间是[0,1]，也就是说，两事件可能完全相同，也可能毫无关系。0(或低于某一定义的阈值)表示两个事件完全不相似，1 表示两个事件完全相似，即是同一个事件。

为更为准确地测算文本事件中的相似度，考虑文本描述前后语句时会出现省略成份的现象，从而导致文本中的部分事件的要素不能清楚地表达，如“一辆宁夏半挂车与同向等候放行的一辆小型客车相撞并起火燃烧。”前后出现 3 个事件触发词“相撞”、“起火”、“燃烧”，但文本仅完整描述事件“相撞”的要素，两事件“起火”、“燃烧”事件并未完整描述它们的要素，仅描述了动作要素。为此，先将要素描述不完整的事件按

事件的定义对进行事件要素的补全。并对事件要素对应的词汇结合同义词林判断事件间对应要素的相似性，如果是同义词(死伤与伤亡)，就认为相似度定义为1；如果存在相互关系(火灾与起火)就似为相近，其相似度定义为0.5等。根据事件要素在文本的地位，将 w_k 设定为 $w_1=0.5, w_2=0.3, w_{3,4,5,6}=0.1$。通过实验观察，当事件相似度 $SIM(e_i,e_j)\geqslant 0.7$时，可以认为$(e_i,e_j)$是相似的。

定义5(事件网络) 事件网络是指一组包含一系列事件结点及相连边的有向图的集合，结点表示事件，边表示事件间的关系。事件间存在关系的采用单向边表示，事件相似度高的事件间用有向双向边表示。形式化表示为：

$$GRE=[Events,Ls,W] \tag{3}$$

$$Events:\{e_1,e_2,e_3,\cdots,e_n\} \tag{4}$$

$$Ls:\{(e_1,e_2,l(e_1,e_2)),(e_1,e_3,l(e_1,e_3))\cdots(e_i,e_j,l(e_i,e_j))\} \tag{5}$$

$$W=l(e_i,e_j),W\in[0,1] \tag{6}$$

其中，GRE 是事件网络。

$Events$ 表示事件集合，节点集合 $N=\{e_1,e_2,\cdots,e_n\}$ 中的每个节点 e_i 代表一个事件特征，n 为整个图结构的节点个数。

Ls 表示事件间的关系集合，在有向边的集合 $E=\{\cdots,l_{ij},\cdots\}$ 中，每条有向边 $l_{ij}(i,j=1,2,\cdots,n,且\ i\neq j)$ 代表两个邻接节点 e_i 和 e_j 对应的事件间的各类关系。

W 表示单位事件同其它事件的链接关系度，即 $l(e_i,e_j)$ 的值，用区间$[0,1]$之间的值来表示，且 $\sum_{i=1}^{n} l(e_i,e_j)=1$。

3 事件网络的构建

为构建事件间的关系，先在已标注语料文本 D 中抽取出事件、事件要素，根据构建的事件本体抽取事件间的关系，得到文本 D 的事件集合 $E=\{e_1,e_2,\cdots,e_i,e_j,\cdots,e_n\}$；并在此基础上构建事件网络，构建步骤如下。

①初始化节点集合 $N_d=\{\}$、有向边集合 $E_d=\{\}$；

②依次将文本 D 的事件集合 $E(D)=\{e_1,e_2,\cdots,e_i,e_j,\cdots,e_n\}$ 中的单位事件映射至事件网络图结构中的节点，得到节点集合 $N_d=\{n_1,n_2,\cdots,n_i,n_j,\cdots,n_k\}$；

③在节点集合 N_d 中取节点 n_i 作为事件网络的任意节点，并在节点集合 N_d 中依次查找与 n_i 相关联的节点 n_j，节点 n_i 和 n_j 间具有组成、因果及跟随关系的则添加一条有向关系边 $\xrightarrow[e_{ij}]{}$，对发生伴随关系的则添加有向关系边 $\xleftrightarrow[e_{ij}]{}$；

④在节点集合 N_d 中任取节点 n_i，依次遍历集合 N_d 中其它节点 n_j，计算它们对应的事件特征 e_i 和 e_j 的相似度，如果相似度大于等于阈值(这里阈值设定为0.7)，则在 n_i 和 n_j 之间添加两条相向的有向边 $\xleftrightarrow[e_{ij}]{}$；

⑤根据③和④可以从而得到有向图集合 $E(D)=\{\cdots,e_{ij},\cdots\}$，从而得到文本 D 的事件网络有向图。

下面列举事件网络的实例说明事件间的关系，文本摘自 http://news.ifeng.com/mainland/detail_2014_03/25/35110264_0.shtml，即“包茂高速黔江段发生一起交通事故”。文本包含4个段落，8个句子，28个事件。使用CEC标注工具对文本 D 进行事件标注，并利用本文的方法得到文本 D 的事件网络有向图，包含28个节点，88条有向边。通过NetDraw将文本可视化，每个节点使用事件及其动作要素进行标识，通过节点间带箭头的线表示事件间的关系，如图2所示。

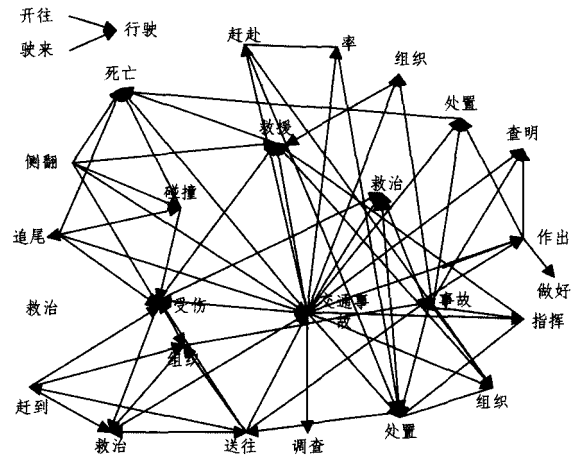

图2 事件网络有向图

4 基于事件主题抽取形成摘要

为获取文本摘要，在对文本中的事件的重要度进行排序后需将文本事件全部串联起来，最终生成摘要。由于文本对整个事件的描述时往往会以相似的语言描述同一事件，因此在获取事件的重要度时，需先根据文本中事件的关系及事件的相似度构建事件网络有向图，再分别计算出事件网络有向图中各节点的重要程度并排序，将重要程度最高的事件称为主题事件。

图3为抽象化的事件网络表示模型。

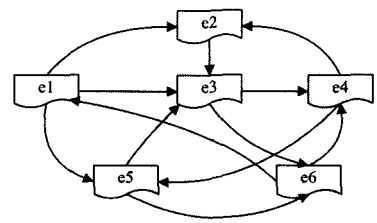

图3 事件网络模型

从图3可以看出，事件与事件间存在一条或多条有向边相互连接，若某节点存在 K 条边链出，N 条边链入，表示当前事件的发生影响 K 个其它事件，同受 N 个其它事件的影响。因此我们可将节点间的链入与链出边通过矩阵的方式来表示事件间的关系。

$$W=\begin{bmatrix} l(e_1,e_1) & l(e_1,e_2) & \cdots & l(e_1,e_n) \\ l(e_2,e_1) & \ddots & & l(e_2,e_n) \\ \vdots & & l(e_i,e_j) & \vdots \\ l(e_n,e_1) & l(e_n,e_2) & \cdots & l(e_n,e_n) \end{bmatrix}$$

其中，$l(e_i,e_j)$表示节点 e_i 指向 e_j 方向的一条边。如果该边存在，则 $\sum_{i=1}^{n} l(e_i,e_j)=1$；否则 $l(e_i,e_j)=0$。

事件的重要程度采用经典 PAGERANK[15] 算法对基于事件相似图的事件进行排序，图中每个节点的度计算公式为：

$$R(e_i)=d\sum_{i\in L(e_i)} R(e_j)/L(e_j)+(1-d)/n, d\in[0,1] \tag{7}$$

其中，e_i,e_j 是事件图中的任意节点，$R(e_i)$表示事件 e_i 的重要度，$R(e_j)$表示事件 e_j 的重要度。$L(e_i)$表示连接线指向 e_i 的集合(导致 e_i 出现的事件总数)，$L(e_j)$表示 e_j 连接线指向别的事件 $e_k(k\in n)$的个数(由 e_j 导致 e_k 的事件总数)。n 是图中节点数(相关联的事件个数)；d 是参数，为一个衰减因子，又称为阻尼系数$[0,1]$，通常取 $d=0.85$。

主题事件排序是生成摘要的重要环节。如果顺序不当，会降低摘要本身的质量和可靠性。因此主题事件排序应按照事件的重要程度及发展过程进行排序。应在重要程度排序基础上按照事件的发展过程进行排序，对于无法比较时间，但属于同一文档且重要程度相同的主题事件按照其出现的先后顺序排序。最后逐步删除排序好的语句集合中对信息贡献最小的句子，直到剩余的句子长度之和达到目标文摘长度。

5 实验结果与分析

5.1 实验数据及评价性能分析

在实验中，本文从 CEC 语料库的 5 个事件类中各随机抽取共 203 篇文本语料，经过分句、词性标注、事件关系项抽取等预处理后，每类文档集的事件项数目以及具有关系的事件项数目的统计如表 1 所列。

表 1 CEC 语料数据统计表(个)

实验语料	文档数	事件数	句子数	带关系事件数	时间短语数
交通事故	54	837	265	1614	203
地震	45	704	292	1208	217
火灾	31	531	260	962	199
食物中毒	43	191	288	322	214
恐怖袭击	30	490	249	880	218
总计	203	2753	1354	4986	1051

目前，自动摘要的评价方法通常采用内部评价(Intrinsic)和外部评价(Extrinsic)。两种评价方法都有各自的优势和劣势，其中内部评价方法简单、容易实现，但主观性太强；外部评价方法较为客观，适用于大规模地对多个摘要系统进行综合评价，但资源消耗大，且评价具有一定局限性。

自动摘要的本质是信息的抽取和压缩，主要采用召回率 R(Recall)、准确率 P(Precision)和调和平均值 F(F-Measure) 3 个指标对自动摘要系统进行内部评价。

摘要召回率反映摘要对原文主题信息的覆盖程度，是对评价摘要质量的一个重要标准。摘要召回率 R 定义为：

$$R=|x\cap y|/|x|*100\%$$

摘要准确率反映摘要表现原文主题信息的准确程度。摘要准确率 P 定义为：

$$P=|x\cap y|/|y|*100\%$$

F 值即为准确率和召回率的调和平均值

$$F=P*R*2/(P+R)$$

5.2 实验结果

为了验证本文自动摘要方法的有效性，选取近期国内外的文档自动摘要研究方法同本文方法进行实验对比。

方法一：首先，以单句作为事件的基本抽取单位，通过二元分类器辨析出事件句和非事件句；然后，通过对事件句聚类，得到同一主题文档集中所包含的不同事件集合，完成事件抽取。

方法二：首先利用滑窗方法抽取主题词，构建空间向量并生成无向图；然后基于向量空间模型计算边权重；最后利用文档句相似度矩阵的权重模型对文档句权重进行建模与计算，依据压缩比得到文档的主题句，完成事件抽取。

方法三：提取单个句子局部属性和句子间的全局属性。句子局部可以被认为是在每个句子的意义的词群，而全局属性可以看成所有文档中的句子之间的关系。对这两个属性组合进行排名和提取句子。

实验中，先用本文方法对每个实验语料生成一个摘要，将本文生成的摘要与方法一、二、三生成的摘要对比，计算上面 3 个指标的值。实验结果如表 2 所列。

表 2 本文实验结果与其他方法研究结果的比较

文章题目	实验语料数据	实验结果		
		P	R	F
基于关键词抽取的自动文摘算法[16]	1998 年《人民日报》中选择了 100 篇文章	0.62	0.51	0.56
基于无向图构建策略的主题句抽取[17]	27 篇新闻类和 20 篇文学类文章，20 篇科技论文，共 67 篇文档	0.57	0.54	0.55
Jaruskulchai[18]	从 Ziff-Davis 语料库中选取了 10 篇文章	0.60	0.44	0.50
本文方法	从 CEC 语料库中 5 个事件类中各随机抽取 20 篇，共 100 篇文本语料	0.65	0.58	0.61

由表 2 中不同方法下不同语料生成摘要对比可看出，本文提出的方法在生成摘要的召回率、准确率和调和平均值上略高于其它类别(也可能是本实验的语料多数局限于突发事件类)。

结束语 针对含有大量事件的突发事件的叙事类文档，本文根据事件间的相互关系构建事件关系有向图网络，带有向图的关系网络不仅清晰地理解事件发展趋势，还能根据事件网络有向图勾画出事件之间的关联性，运用经典 PAGERANK 计算各节点的相对重要度，从而得到各子事件在整个事件的重要程度，并在事件的重要程度排序基础上按照事件的发展先后进行排序。最后逐步删除排序好的语句集合中对信息贡献最小的句子，直到剩余的句子长度之和达到目标文摘长度。实验证明这种方法能更好地概括出文本的主要内容。

参 考 文 献

[1] 胡侠，等. 自动文本摘要技术综述[J]. 情报杂志，2010，29(8)：144-147

[2] Lin C Y. Training a Selection Function for Extraction[C]//Proceeding of the Eighteenth Annual International ACM Conference on Information and Knowledge Management (CIKM). 1999：55-62

[3] Jiang Chang-jin，Peng Hong，Ma Qian-li，et al. Automatic Summarization for Chinese Text Based on Combined Words Recognition and Paragraph Clustering[C]// Proceedings of the 2010 Third International Symposium on Intelligent Information Technology and Security Informatics(IITSI'10). 2010：591-594

[4] Z Pei-ying，L Cun-he. Automatic text summarization based on sentences clustering and extraction[C]// Proceedings of the 2009 2nd International Conference on Computer Science and Information Technology (ICCSIT 2009). 2009：167-170

[5] Chandra M，Gupta V，Paul S K. A Statistical approach for Automatic Text Summarization by Extraction[C]//2011 International Conference on Communication Systems and Network Technologies(CSNT 2011). 2011：268-271

[6] Erkan G，Radev D R. LexRank：Graph-based lexical centrality as salience in text summarization[J]. Journal of Artificial Intelligence Research，2004(22)：457-479

[7] Zwaan R A，Radvansky G A. Situation models in language comprehension and memory[J]. Psychological bulletin，1998，123(2)：162-185

(下转第 223 页)

响，并且文化程度与生育年龄、生育世代间隔差成反比关系。

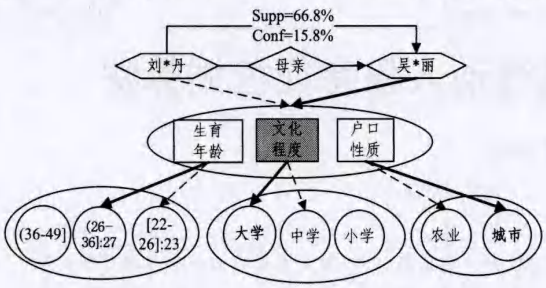

图 9　关联规则实例简化表示

本文所提方法能够有效地实现基于 Vis-Meta 图的关联规则知识表示，并且规则信息表示形式简洁明确。与基于表、基于矩阵、基于 TwoKey 图、基于概念图以及基于平行坐标的方法[8,13,14]相比具有以下优点：能够实现一对一、一对多、多对一以及多对多的多模式关联规则实例化表示，表现力强；能充分展示规则的属性信息和相关领域知识，帮助用户获取更为精确的规则信息以及隐含的知识；表示结果布局清晰，具有较强的可解释性，便于普通用户理解。

结束语　本文给出了 Vis-Meta 图的概念，引入概念关系实现了 Vis-Meta 图应用于关联规则知识表示，所设计的方法能够明确展示领域知识与规则项，不仅能够进行多模式规则实例的对比，而且能够通过展示领域知识辅助规则分析，层次对称结构便于将表示符号转换为自然语言，降低普通用户的感知负担，在一定程度上实现了知识共享。通过将所提算法应用于某省全员人口数据规则，表明该方法具有良好的展示效果，用户在规则知识的基础上发现了新的潜在的知识，为决策提供了有力支撑。

参 考 文 献

[1] Basu A, Blanning R. Metagraphs and Their Applications [M]. Berlin: Springer-Verlag, 2006: 1-11, 77-115

[2] Gaur D. Metagraph a New Hierarchical Data Structured As a Dicision Tree[J]. The Journal of Computer Science and Information Technology, 2007, 6(1): 1-5

[3] Gaur D, Shastri A, Biswas R. Metagraph-Based Substructure Pattern mining[C] // International Conference on Advanced Computer Theory and Engineering, 2008. (ICACTE'08). IEEE, 2008: 865-869

[4] Hu Zen-jun, Mellor J, Wu Jie, et al. Towards zoomable multidimensional maps of the cell[J]. Nature biotechnology, 2007, 25(5): 547-554

[5] Dashore P, Jain S, Dashore S R. Fuzzy Metagraph and Rule Based System for Decision Making in Share Market[J]. International Journal of Computer Applications, 2010, 6(2): 10-13

[6] Dashore P, Jain S. Fuzzy Rule Based Expert System to Represent Uncertain Knowledge of E-commerce[J]. International Journal of Computer Theory and Engineering, 2010, 2: 882-886

[7] Mukherjee A, Sen A K, Bagchi A. The representation, analysis and verification of business processes: a metagraph-based approach[J]. Information Technology and Management, 2007, 8(1): 65-81

[8] 郭晓波，赵书良，刘军丹，等. 基于概念图的关联规则知识表示[J]. 计算机科学，2013，40(8)：261-265

[9] 谭政华，胡光锐，任晓林. 模糊元图及其特性分析[J]. 计算机研究与发展，2000(3)：272-277

[10] Velazquez-Garcia E, Lopez-Arevalo I, Sosa-Sosa V. Distributed Computing and Artificial Intelligence[M]. Springer Berlin Heidelberg, 2012: 469-476

[11] Jain P D S K. Fuzzy rule based system and metagraph for risk management in electronic banking activities[J]. International Journal of Engineering and Technology, 2009, 1(1): 1793-8236

[12] Tan Z H. Fuzzy metagraph and its combination with the indexing approach in rule-based systems[J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(6): 829-841

[13] Bruzzese D, Davino C. Visual mining of association rules[C] // Visual Data Mining: Theory, Techniques and Tools for Visual Analytics, LNAI 6208. Berlin: Springer-Verlag, 2008: 103-122

[14] Liu Gui-mei, Suchitra A, Zhang Hao-jun, et al. AssocExplorer: an association rule visualization system for exploratory data analysis[C] // Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. New York: ACM, 2012: 1536-1539

(上接第 213 页)

[8] ACE (Automatic Contente Extraction) Chinese Annotation Guidelines for Events[R]. National Institute of Standards and Technology, 2005

[9] 刘茂福，李文捷，姬东鸿. 基于事件项语义图聚类的多文档摘要方法[J]. 中文信息学报，2010，24(5)：77-84

[10] 韩永峰. 基于事件抽取的网络新闻多文档自动摘要[J]. 中文信息学报，2012，26(1)：58-66

[11] Ge Shu-zhi, Zhang Zheng-chen, He Hong-shen. Weighted Graph Model Based Sentence Clustering and Ranking for Document Summarization[C] // 4th International Conference on Interaction Sciences (ICIS). 2011: 90-95

[12] Thwaites P. Causal identifiability via Chain Event Graphs[J]. Artificial Intelligence, 2013(195): 291-315

[13] Zhong Zhao-man, Liu Zong-tian. Ranking Events Based on Event Relation Graph for a Single Document [J]. Information Technology Journal, 2010, 9(1): 174-178

[14] 刘宗田，黄美丽，周文，等. 面向事件的本体研究[J]. 计算机科学，2009，36(11)：189-192，199

[15] Page L, Brin S, Motwani R, et al. The Pagerank citation ranking: Bringing order to the Web, Technical report[J]. Stanford University, 1998

[16] 蒋效宇. 基于关键词抽取的自动文摘算法[J]. 计算机工程，2012，38(3)：183-186

[17] 葛斌，李芳芳，李阜，等. 基于无向图构建策略的主题句抽取[J]. 计算机科学，2011，38(5)：181-185

[18] Jaruskulchai C, Kruengkrai C. Generic text summarization using local and global properties of sentences[C] // Proceedings of the IEEE/WIC International Conference on Web Intelligence. Piscataway, USA: IEEE Press, 2003: 201-206