

# 基于中文关键词提取的预案智能匹配方案

金保华, 林青, 吴怀广

(郑州轻工业学院 计算机与通信工程学院, 河南 郑州 450001)

**摘要:**针对当下在公共社会领域中突发事件应急预案结构复杂、难以为决策者进行信息筛选和整合的热点问题,提出了一种基于中文关键词提取的预案智能匹配方案.在运用框架理论解决预案结构存储的基础上,对存储的预案文本进行预处理分词;选用词语的词频、词性以及存在于预案框架中的位置作为提取出来的中文关键词特征值;利用均方误差最小的线性估计法训练关键词权重的调节因子,能够有效提高关键词提取的准确率.实验结果表明,在预案智能匹配方案中运用此种中文关键词提取算法比传统的FIFA算法效果更佳.

**关键词:**应急预案;中文关键词提取;智能匹配;矢量空间模型

**中图分类号:**TP391 **文献标志码:**A **DOI:**10.3969/j.issn.2095-476X.2013.02.018

## Intelligent match scheme of emergency plan based on Chinese keywords extraction

JIN Bao-hua, LIN Qing, WU Huai-guang

(College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450001, China)

**Abstract:** The complication of emergency plan in public social field make plans hard to manage. And it also make decision maker hard to decide and integrate information. An intelligent match scheme of emergency plan based on Chinese keywords extraction was introduced. Based on the solving structure storage of emergency plan using framework structure theory, the participle of storage emergency plan was processed, the frequency, part of speech, position of framework were selected as keywords' characteristic values, least error of mean square method was used for training keywords weight regulatory factors to realize intelligent match of emergency plans more exact and efficient. The experimental results showed that the presented method is more promising in intelligent match method of emergency plan than tradition FIFA method.

**Key words:** emergency plan; Chinese keywords extraction; intelligent match; vector space model

## 0 引言

应急预案是突发事件中组织救援抢险的重要手段,同时也是反映政府部门快速反应能力的一把

标尺.目前,国内的应急预案管理系统主要以对预案进行查询、添加和修改等日常操作为主,忽略了预案资源的重用性.对预案资源的高层次利用包括运用匹配、推理等手段对预案实行智能化管理.由

收稿日期:2012-12-04

基金项目:郑州轻工业学院博士基金项目(2011BSJJ015)

作者简介:金保华(1966—),男,河南省郑州市人,郑州轻工业学院副教授,博士,主要研究方向为人工智能、计算机辅助决策系统.

于对应急预案的重要性认识不足,加之资源的信息量巨大,使得进行匹配时耗费的时间和空间也比较大。

本文将引入框架理论<sup>[1]</sup>,分析中文关键词的特征值与关键词权重调节因子的训练方法,以建立矢量空间模型,更高效更准确地实现预案间的智能匹配。

## 1 预案智能匹配的设计

### 1.1 框架理论在预案智能匹配中的应用

预案智能匹配模块是智能管理系统的核心成分。在预案智能匹配模块中,引入框架理论解决预案的智能化存储问题,以更好地应对预案临机决策问题。同时,这种存储方式也实现了传统文本预案的电子化,数据与程序相互独立,实现了网络共享和数据同步,可对相关的信息进行动态调整,具有动态双向查询功能,减少了大量人力。

框架理论是1975年由Minsky首次提出的。框架理论的核心是以框架这种形式来表示知识。框架的顶层是固定的,下层通过对若干槽(slot)与若干侧面(facet)的取值范围、求值方法的说明形成框架系统,完成资源知识结构的表达<sup>[2]</sup>。

框架定制不仅具有很好的扩展性能,而且可以通过利用槽与侧面中的信息去调用相应的模型,为智能预案的实现预留出接口。

经过一系列的研究,笔者发现在预案框架中,在预案构成的内容上,“总则”、“组织机构及职责”、“应急响应”是预案中必不可少的元素,即这个预案的核心成分。因此,这3个槽可以作为预案框架中的基本槽。而“总则”槽中的基本元素“适用范围”,“组织机构及职责”槽中的基本元素有“应急指挥体系以及职责”侧面,因此,这2个侧面可以作为响应槽中的基本侧面。基于基本预案评估所得出的匹配不仅仅是最原始的匹配值,也是整个综合预案中核心成分中最重要基础的匹配,因此,这些从继承父类所得到的对象属性对于一个预案的智能匹配模式有着非常重要的意义。

### 1.2 基于中文文本关键词提取的预案智能匹配总体设计

随着国家应急救援平台的发展,预案信息量日趋庞大<sup>[3]</sup>,同时对预案信息实时更新、分类管理提出了更高的要求。因此,在预案的智能匹配模式上,

动态地存储大量的文本数据是不现实的。为了进一步降低网络运行环境上和系统上的内耗,同时保证应急救援防范工作中预案信息的准确体现,本文提出了基于候选词权重的中文文本关键词提取方法。该方法从页面读取框架信息进行相应的中文文本关键词提取,使用均方误差最小的线性估计法训练调节因子自学习,存储并处理相应字段信息进行匹配。

在预案智能管理系统中,智能匹配模块提供预案定制、预案初始化、参数调整、智能匹配业务进行匹配工作。预案的智能匹配流程可用图1来表示。

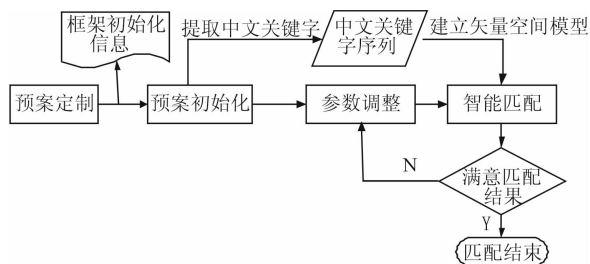


图1 预案的智能匹配流程图

预案定制,指的是预案模板的定制。除了“总则”、“应急组织体系”、“应急响应”这3个必要的槽之外,用户依照预案要素为该预案加入槽的信息。以《河南省盐业市场突发事件应急预案》为例,用户需要添“后期处理”槽和“保障措施”槽,同样地,为“总则”槽加入“编制目的”“应急工作原则”与“编制依据”这3个侧面,为“应急组织体系”槽内加入“应急组织机构”侧面。至此,整个预案的定制就完成了。

预案初始化,指的是将相应的文本信息通过文本导入的方式,通过生成的定制模板为预案进行初始化操作,通过表单的提交转入后台中文文本关键词提取并存入 Oracle 数据库中。

参数调整是为后期的智能匹配提供服务的。依照预案的28个要素的相应信息,用户将可侧重地调整预案框架中的槽与侧面的参数,从而为用户提供专业定制级的匹配服务。值得注意的是,在响应槽中的基本槽、基本侧面时使用的是模糊匹配,以增加匹配的实用性。

通过参数调整操作,依照预案的28个要素的相应信息,有侧重地勾选需要匹配到的槽或侧面,预案智能匹配问题,实则是预案与预案需匹配的框架信息中需匹配的文本与文本的相似度问题。记预案

$x$  与预案  $y$  的相似度为  $\text{sim}(plan_x, plan_y)$ , 因为 1 个预案需匹配的部分也是由槽与侧面组成的, 因此, 运用矢量空间模型来计算  $plan_x$  与  $plan_y$  之间的相似度. 将槽(此时槽中只有槽值)与侧面中的文本信息使用上述关键词提取得到的词组, 分别记为  $(ws_1, ws_2, \dots, ws_m)$  与  $(wf_1, wf_2, \dots, wf_m)$ . 则

$$\text{sim}(plan_x, plan_y) =$$

$$\begin{cases} \sum_{i=1}^n \text{sim}(slot_{xs1}, slot_{ys1}) & \text{若需匹配不存在侧面的槽} \\ \sum_{i=1}^n \text{sim}(slot_{xs2}, slot_{ys2}) & \text{若需匹配存在侧面的槽} \end{cases}$$

$$\text{sim}(slot_{xs2}, slot_{ys2}) = \sum_{i=1}^n \text{sim}(facet_{xi}, facet_{yi})$$

利用矢量空间模型<sup>[4]</sup>, 可知

$$\text{sim}(slot_{xs1}, slot_{ys1}) = \cos(slot_{xs1}, slot_{ys1}) =$$

$$\frac{slot_{xs1} \times slot_{ys1}}{\|slot_{xs1}\| \times \|slot_{ys1}\|} = - \frac{\sum_{i=1}^n (w_{xsi} \times w_{ysi})}{\sqrt{\sum_{i=1}^n w_{xsi}^2} \cdot \sqrt{\sum_{i=1}^n w_{ysi}^2}}$$

同理可知

$$\text{sim}(facet_{xi}, facet_{yi}) = \cos(facet_{xi}, facet_{yi}) =$$

$$\frac{facet_{xi} \times facet_{yi}}{\|facet_{xi}\| \times \|facet_{yi}\|} = - \frac{\sum_{i=1}^n (w_{xfi} \times w_{yfi})}{\sqrt{\sum_{i=1}^n w_{xfi}^2} \cdot \sqrt{\sum_{i=1}^n w_{yfi}^2}}$$

预案匹配模块最终以相似度按正序列出匹配结果. 通过上述方法, 用户可以灵活地按自身决策方案的需要最大程度地得到需要匹配的预案, 从而形成合理的决策方案, 实现预案价值的最大化.

## 2 中文文本关键词提取算法

### 2.1 关键词的预处理

该系统在分词上采用的是 imdict-chinese-analyzer 开源分词工具. 运用分词技术处理框架信息文本, 即在分词的基础上, 进行相应的处理<sup>[5]</sup>. 除此之外, 如何自动识别词的边界, 从而将汉字序列分为正确的词串的中文分词问题, 将是实现中文信息处理的基础问题<sup>[6]</sup>.

在分词的基础上, 进行相应的处理, 指的是框架信息文本中的预处理, 包括去除停用词以及进行词性过滤. 所谓停用词, 是指那些不能反映主题的功能词, 它们不但不能反映文献的主题, 而且还会对主题词的抽取造成干扰, 有必要将其滤除<sup>[7]</sup>. 停用词一般包括标点符号、虚词, 基于预案的特殊性,

还包含一些对需要匹配的主题反映程度不够强烈的词语. 像文本中的标点和虚词是很好理解的, 因为这些词语词频较高, 但对反映的主题没有任何帮助, 反而带来很高的干扰性, 影响分词的效果. 根据反复测试预案文本的过程得到的动词停用表, 是出于进一步提高提取文本特征的准确性的考虑.

### 2.2 关键词提取算法的特征项

在预案初始化操作中用户导入的文本信息经过分词与词性过滤后得到词集, 所面临的问题就是选择关键词的问题. 选择以下 3 个因子作为关键词的特征项.

1) 词频: 它是有效反映该词是否能反映主题的标准之一. 运用以下公式, 作为词语  $i$  的词频因子  $freq_i$ .

$$freq_i = f_i(f_i + 1) \quad \textcircled{1}$$

其中,  $f_i$  表示是词语  $i$  在用户导入到定制好的预案框架中的文本信息中的词频. ①式是非线性函数, 即词频因子随着词频的增加逐渐上升, 当词频足够大时, 函数向 1 收敛.

2) 词性: 预案作为一个具有行政意义的方案<sup>[8-9]</sup>, 其公文性较强, 笔者认为, 名词在预案中较能反映主题. 运用以下公式, 计算词语  $i$  经过分词与词性过滤后得到的词性因子  $type_i$ .

$$type_i = (w_i + 1)(w_i - 1)$$

其中,  $w_i$  的取值为

$$w_i = \begin{cases} 3.0 & \text{名词} \\ 2.0 & \text{动词} \\ 1.0 & \text{其他词语} \end{cases}$$

3) 词语在定制的模板的位置. 根据预案间的相似度的计算可知, 词语在基本槽与基本侧面上的位置具有重要的意义. 运用以下公式计算词语  $i$  的位置因子  $pos_i$ .

$$pos_i = \begin{cases} 0.8 & \text{当词语 } i \text{ 出现在基本槽的基本侧面中} \\ 0.6 & \text{当词语 } i \text{ 出现在基本槽却不在基本侧面中} \\ 0.4 & \text{当词语 } i \text{ 不出现在基本槽中} \end{cases}$$

其中,  $pos_i$  为词语  $i$  在预案导入到定制好的预案框架中文本信息的位置因子.

### 2.3 关键词权重计算

所有的关键词权重计算都是基于上文所述的词频因子、词性因子与位置因子这 3 个因子, 词语  $i$  计算公式如下:

$$weight_i = \alpha \times freq_i + \beta \times type_i + \gamma \times pos_i$$

其中,  $weight_i$  为预案框架文本框中候选词  $i$  的权重;

$freq_i$  为词频因子;  $type_i$  为词性因子;  $pos_i$  为位置因子;  $\alpha, \beta, \gamma$  为 3 个因子的调节因子。

关键词的权重是指关键词能够反映预案主题的程度。为了能够更有效地计算出关键词的权重,可以通过均方误差最小的线性估计法训练调节因子。具体方法如下。

1) 设定该预案框架信息中文本的训练词语排序  $v_i$ 。

2) 将第 1 个训练文本依照初始调节因子的值计算出词语的权重并排序  $v_1$ , 并将其中词语  $i$  得到的  $freq_i, type_i, pos_i$  分别记为  $(A_i, B_i, C_i)$ , 因此从词语 1 到词语  $n$  所得到的词语序列为

$$A = \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ \vdots & \vdots & \vdots \\ A_n & B_n & C_n \end{bmatrix}$$

3) 根据  $v_1$  反解出关于  $(\alpha \beta \gamma)^T$  的解集  $F$ , 记为  $\eta = (\alpha_i \beta_i \gamma_i)^T$ 。

4) 根据  $\eta$  计算出

$$\xi(\eta) = \begin{pmatrix} A_1 \times \alpha + B_1 \times \beta + C_1 \times \gamma \\ A_2 \times \alpha + B_2 \times \beta + C_2 \times \gamma \\ \vdots \\ A_n \times \alpha + B_n \times \beta + C_n \times \gamma \end{pmatrix}$$

5) 将调节因子矩阵  $(\alpha \beta \gamma)^T$  记为  $b$ 。

6) 关于  $\xi$  的  $\eta$  估计  $\hat{\xi}(\eta) = A\eta + b$ , 为使得  $E\{\|\xi - \hat{\xi}\|^2\} = E\{\|\xi - (A\eta + b)\|^2\}$  达到最小值, 这时  $\hat{\xi} = A\eta + b$  称作为最佳线性估值。其中  $\xi$  是  $n \times 1$  的矢量,  $\eta$  是  $m \times 1$  的矢量,  $A$  是  $n \times m$  的矩阵,  $b$  是  $n \times 1$  的矢量。此时上述关于  $\xi$  的  $\eta$  最小均方误差估计  $\hat{\xi}(\eta) = E\{\xi/\eta\}$  问题转化为需要找到一个关于  $\eta$  的线性函数作为最佳估计。

7) 利用正交性原理, 即如果用复随机变量  $\eta_1, \eta_2, \dots, \eta_m$  的线性组合  $\hat{\xi} = a_1 \times \eta_1 + a_2 \times \eta_2 + \dots + a_m \times \eta_m$  来估计复随机变量  $\xi$ , 其估计误差  $e = \xi - \hat{\xi}$ 。若  $e$  和  $\eta_1, \eta_2, \dots, \eta_m$  正交, 则此估计是最小均方误差线性估计。利用  $E\{e \cdot \overline{\eta_j}\} = E\{[\xi - (a_1 \eta_1 + a_2 \eta_2 + \dots + a_m \eta_m)] \cdot \overline{\eta_j}\} = 0$  的满足正交性的线性估计来计算得到  $b = E\xi - A_j \times E\eta$ 。并利用  $b$  训练得到矢量  $\eta' = (\alpha'_i \beta'_i \gamma'_i)^T$ 。

8) 重复 3) 到 7), 得到第  $k$  次计算得到的  $\eta^k = (\alpha^k_i \beta^k_i \gamma^k_i)^T$ 。

9) 对  $s$  个文本重复 2) 到 8) 的过程, 将得到的

$\eta^{k'}, \eta^{k''}, \dots, \eta^{k_s}$  则可以得到

$$\eta_s = (1/s)(\eta^{k'} + \eta^{k''} + \dots + \eta^{k_s}) \quad (2)$$

$\eta_s$  为经过训练文本调整之后得到的关于  $\alpha, \beta, \gamma$  的向量。

10) 将  $\eta_s$  根据 ② 式进行计算, 即可得到词语  $n$  的权重, 并默认取框架信息输入文本中权重值最大的前 5 个形成数组存入数据库字段中。

## 3 实验结果与分析

### 3.1 实验语料

本实验所使用的语料为应急平台信息技术河南省工程实验室的预案语料库, 其中的预案文档严格按照预案的分类方法, 即按预案类型主要分为自然灾害、事故灾难、公共卫生和社会安全 4 个类作为语料<sup>[10-11]</sup>。从各个类中选择一定数量的文档, 并从文档中节选适当的章节作为定义的预案, 定制部分的文本作为训练集, 而将上文所提及的调查系统网络与信息安全应急预案作为测试集。

对于这些训练集与测试集, 其关键词都是人工赋予的, 其中训练集的关键词数目的选取依据为每个文本的长度。

### 3.2 评价方法

对于预案匹配中关键词提取算法的评估方法依靠对调查系统网络与信息安全应急预案测试集中算法提取的关键词与标准关键词作语法上的匹配后, 所得到的中文文本提取关键词的平均查准率  $P$  与召回率  $R$  这 2 个因素来核查算法的效果<sup>[12]</sup>。而此处所提及的标准关键词是由 5 位计算机应用专业的研究生组成的人工标注小组经过反复的互阅所得。其中, 查准率  $P$  与召回率  $R$  的计算方法分别如下:

$$P = \frac{n_c}{n_a} \quad R = \frac{n_c}{n_r}$$

其中,  $n_c$  为自动提取正确的关键词数目,  $n_a$  为自动提取的关键词数目,  $n_r$  为人工赋予的完全的关键词数目。

### 3.3 分析与讨论

图 2—图 4 为对 10 篇, 20 篇, 30 篇, 40 篇, 50 篇文本进行训练后, 分别对测试集的预案提取 5 个关键词、10 个关键词和 15 个关键词的情况。从这 3 幅图可以得到一个共同的结论: 在对大小不同的训练文本提取相同数目的关键词的情况下, 随着训练文本数量的增加, 查准率与召回率均有较大提高; 但

当训练文本增加到一定数目后,查准率与召回率增加情况相对较为平稳.由此可以推断出,经过均方误差最小的线性估计法训练调节因子后,调节因子趋于收敛.

同时,比较图2—图4可以得知,提取的关键词越多,查准率与召回率越高.而当提取的关键词达到一个阈值后,查准率与召回率趋于收敛,由此可知,在训练了此30篇文本平均长度的情况下对调查系统网络与信息安全应急预案提取关键词,适宜采用提取10个左右的关键词来达到提高效率与提取质量的比值.

图5与图6给出了当训练文本为30的时候,从测试集依次提取出5个,10个,15个,20个,25个关键词,此时该算法与FIFA算法<sup>[13]</sup>中的查准率与召回率之间的比较情况.

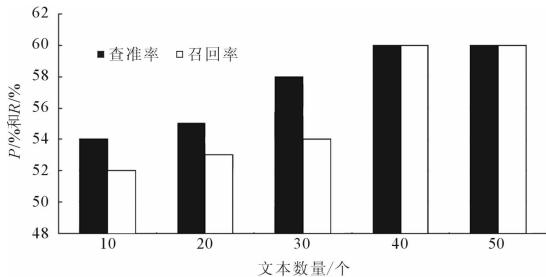


图2 不同训练集大小下的评估因素取值 (提取5个关键词)

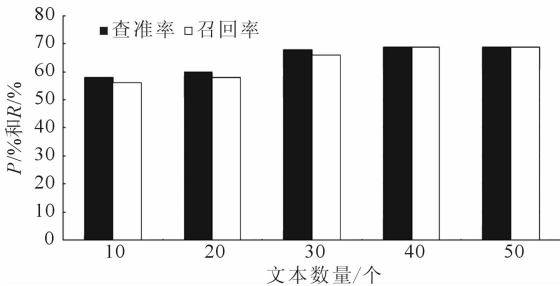


图3 不同训练集大小下的评估因素取值 (提取10个关键词)

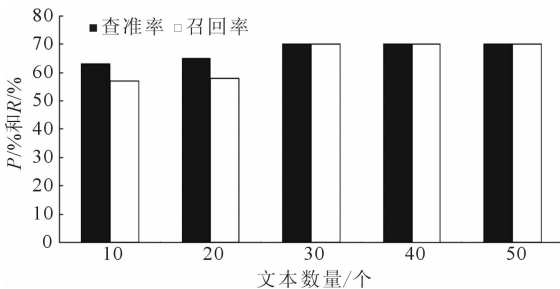


图4 不同训练集大小下的评估因素取值 (提取15个关键词)

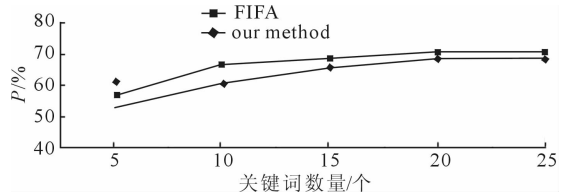


图5 与传统方法相比较的查准率

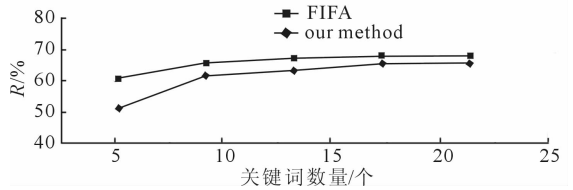


图6 与传统方法相比较的召回率

由图5与图6可知,相对于FIFA算法,采用均方误差最小的线性估计法来训练提取公式,能够有效提高所提取关键词的性能,使得本算法所得到的关键词的查准率在平均高出2%到6%,召回率平均高出4%到6%.

### 4 结论

本文综合分析了预案智能管理中预案的特性,为了能够加强预案资源中各类基础信息的整合,为决策者提供智能匹配方案.本文引入了框架理论,采用词频、词性以及词语在预案框架中的位置作为关键词的权重的特征值;利用均方误差最小的线性估计法训练关键词权重的调节因子,并对分词后关键词进行权重计算;利用矢量空间模型实现预案文本间最大相似度的匹配,完成智能预案匹配的设计;设置参数调整模块用于异类的预案间的匹配,具有良好的扩展性与实用性.

基于中文关键词提取的预案匹配打破了传统的单一的预案匹配模式,可有侧重地选择相应的预案框架进行匹配,用户自主性较大,也为在实战中利用智能预案系统收集反馈信息提供了方便.

### 参考文献:

[1] Stuart Russel. 人工智能——一种现代的方法[M]. 北京:人民邮电出版社,2010.  
 [2] 蔡艳婧,程显毅,潘燕.面向自然语言处理的人工智能框架[J].微电子学与计算机,2011,28(10):173.  
 [3] 井悦.应急预案和应急资源数据库管理系统研究与实现[J].煤炭经济研究,2010,30(7):92.

结果可以看出改进后的 OPTICS 算法识别精度有了一定的提高.

表3 OPTICS 算法聚类后分类结果 %

Application 具体协议	trace1 流/包	trace2 流/包	trace3 流/包
Web	3.05(13.31)	4.90(15.38)	1.89(10.24)
QQ	1.81(6.01)	3.78(11.45)	3.57(16.59)
PPLive	38.53(29.07)	36.32(23.64)	35.16(35.19)
DNS	1.26(1.02)	1.57(6.57)	0.65(1.95)
Bit Torrent	38.31(37.50)	29.02(22.54)	48.97(27.42)
Non-pay load	17.04(8.22)	24.41(9.01)	8.51(5.70)

表4 改进的 OPTICS 算法聚类后分类结果 %

Application 具体协议	trace1 流/包	trace2 流/包	trace3 流/包
Web	3.05(13.31)	4.92(16.08)	2.03(10.45)
QQ	1.91(6.21)	3.76(10.75)	3.42(16.38)
PPLive	38.43(28.87)	36.32(23.64)	35.16(35.19)
DNS	1.26(1.02)	1.57(6.57)	0.65(1.95)
Bit Torrent	38.42(37.71)	29.07(22.68)	49.03(27.56)
Non-pay load	16.93(8.02)	24.36(8.87)	8.45(5.56)

### 4 结论

本文采用的方案在早期用 DPI 技术对数据流进行检测识别,辨识出数据流的业务类型,作为后期聚类个数的指导元素.针对未识别的数据流,通过 OPTICS 算法进行聚类,可以确定加密的无法识别数据流的业务类型.在该方案中后期,聚类算法不受人为因素和初始聚类中心和数据顺序的干扰.同时

该方案前期识别出的业务类别可以作为后期聚类算法准确度的基准,检验该聚类方法的有效性和准确性.

### 参考文献:

[1] Erman J, Arlitt M, Mahanti A. Traffic classification using clustering algorithms [C]//SIGCOMM '06 MineNet Workshop, Pisa: ACM, 2006:11.

[2] 邱密, 阳爱民, 刘永定, 等. 使用贝叶斯学习算法分类网络流量[J]. 计算机工程与应用, 2010, 46(25):78.

[3] 张剑, 钱宗钰, 寿国础. 在线聚类的网络流量识别[J]. 北京邮电大学学报, 2011, 34(1):103.

[4] 钮晓娜, 郭云飞, 张进. 基于数据包抽样的互联网业务流精确分类[J]. 计算机应用研究, 2009, 26(10):3863.

[5] Huang N F, Jai G Y, Chao H C. Early identifying application traffic with application characteristics [C]//Proc of ICC, Beijing: IEEE Conference Publications, 2008:5788.

[6] Moore A W, Papagiannakik. Toward the accurate identification of network applications [C]//Proc of PAM 2005 (LNCS3431), Boston: [s. n. ], 2005:41.

[7] 焦小焦, 钟诚, 杨柳, 等. 基于流量特征和载荷特征的 P2P 流量识别 [J]. 计算机工程与科学, 2010, 32(8):24.

[8] 马永立, 钱宗钰, 寿国础, 等. 机器学习的 Internet 流量分类[J]. 计算机应用, 2010, 30(1):80.

[9] 李卫, 边江, 王盈. 动态网络流量分类研究[J]. 电子科技大学学报, 2007, 36(6):1508.

[10] 马超群, 兰秋军, 陈伟民. 数据挖掘[M]. 北京: 科学出版社, 2007.

(上接第 82 页)

[4] 李学俊. 基于分词算法 &VSM 的文本主观题自动评分算法研究 [J]. 电脑知识与技术: 学术交流, 2011(9):6122.

[5] 金翔宇, 孙正兴, 张福炎. 一种基于词汇链的关键词抽取方法[J]. 中文信息学报, 2001, 15(6):25.

[6] 徐亚娟. 基于公安业务信息的文本挖掘技术研究与实践[D]. 杭州: 浙江大学, 2008.

[7] 郑家恒, 卢娇丽. 关键词抽取方法的研究[J]. 计算机工程, 2005, 31(18):194.

[8] 郝彬. 应急预案行政法律问题研究[D]. 北京: 中国政法大学, 2010.

[9] 栾盈菊. 对政府应急预案编制与实施的思考——以 2008 年南方低温雨雪冰冻灾害为例[J]. 江南社会学院学报, 2008, 10(2):27.

[10] 刘功智, 刘铁民. 重大事故应急预案编制指南[J]. 劳动保护, 2004, 52(4):11.

[11] 金碧芳. 交通紧急事件处理预案仿真评估系统的设计与实现[J]. 计算机应用与软件, 2010, 27(3):157.

[12] 李晓黎, 刘继敏, 史忠植. 概念推理网及其在文本分类中的应用[J]. 计算机研究与发展, 2000, 37(9):1032.

[13] 朱靖波, 姚天顺. 基于 FIFA 算法的文本分类[J]. 中文信息学报, 2002, 16(3):20.