

# 基于关联规则后件扩展的越英跨语言信息检索

黄武锋,何冬蕾,黄名选

(广西财经学院 信息与统计学院,广西 南宁 530003)

**摘要:**针对跨语言信息检索中存在的查询主题漂移问题,提出一种基于完全加权关联规则后件扩展的越英跨语言信息检索模型,给出了模型结构及其各个功能模块,详细阐述了模型的关键技术及其算法。该模型将完全加权模式挖掘技术和用户相关反馈扩展融合应用于越英跨语言信息检索,将越南语查询通过机器翻译系统译为英文并检索为英文文档,提取前列初检文档构建用户相关反馈文档集,采用完全加权关联规则挖掘技术对用户相关反馈文档集挖掘与原查询相关的关联规则,将关联规则后件作为扩展词,并和原查询组合成新查询再次检索英文文档,得到最终检索结果。在 NTCIR-5 CLIR 数据集上的实验结果表明,该模型能减少越英跨语言检索中的查询漂移,提高和改善其检索性能。

**关键词:**完全加权模式挖掘;查询扩展;跨语言信息检索;信息检索

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2019)04-0164-05

doi:10.3969/j.issn.1673-629X.2019.04.033

## Vietnamese-English Cross Language Information Retrieval Model Based on Association Rule Consequent Expansion

HUANG Wu-feng, HE Dong-lei, HUANG Ming-xuan

(School of Information and Statistics, Guangxi University of Finance and Economics, Nanning 530003, China)

**Abstract:** We propose a Vietnamese-English cross language information retrieval model based on all-weighted association rule consequent expansion to solve the problem of query drift existing in cross language information retrieval. The structure of the model and its function modules are given, and the key techniques and algorithms of the model are discussed in detail. This model integrates the techniques of all-weighted pattern mining and user relevance feedback expansion for Vietnamese-English cross language information retrieval, and translates the Vietnamese query into English by machine translation system so as to retrieve English documents, and extracts the top-ranked retrieved documents with the aim of setting up user relevance feedback document collection. The technique of all-weighted association rule mining is used to mine association rules related to the original query in the collection, and the association rule consequents are taken as the expansion terms, and combined with the original query as a new query to retrieve the English documents for the final search result. Experimental results on the NTCIR-5 CLIR data set show that the proposed model can effectively reduce query drift in Vietnamese-English cross language retrieval, and improve its retrieval performance.

**Key words:** all-weighted patterns mining; query expansion; cross language information retrieval; information retrieval

### 1 概述

跨语言信息检索(cross-language information retrieval, CLIR)指的是以一种语言的查询去检索其他语言信息资源的技术,其中表达查询的语言称为源语言,所检索文档的语言称为目标语言。长期以来,跨语言信息检索存在严重的查询主题严重漂移、词不匹配以及查询项翻译歧义和多义性等问题,这些问题一直是制约跨语言信息检索技术发展的瓶颈。近年来,跨语

言信息检索模型研究取得了丰富的研究成果,主要集中在基于查询翻译优化的和基于查询扩展的跨语言信息检索等方面。

基于查询翻译优化的跨语言信息检索针对源语言查询翻译过程中如何得到最优的、最准确的目标语言查询项开展研究。其典型的工作有:Geraldo等<sup>[1]</sup>提出了基于关联规则挖掘的西班牙语-英语跨语言信息检索方法;姚寒冰等<sup>[2]</sup>针对跨境电商涉及不同语言的供

收稿日期:2018-05-05

修回日期:2018-09-13

网络出版时间:2018-12-20

基金项目:国家自然科学基金(61762006,61262028)

作者简介:黄武锋(1975-),男,硕士,讲师,研究方向为数据挖掘、电子商务;黄名选,硕士,教授,通信作者,研究方向为数据挖掘、信息检索和机器学习。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.TP.20181220.1049.042.html>

需双方信息的跨语言检索需求,提出了基于自然语言的跨语言协同机器翻译的信息检索算法;吴丹等<sup>[3]</sup>提出基于相关反馈的跨语言信息检索查询翻译优化技术;Giang等<sup>[4]</sup>提出基于词典翻译的越英跨语言信息检索方法,在此基础上提出了基于伪相关反馈的越英跨语言检索模型<sup>[5]</sup>,有效提升了检索性能;Debasis等<sup>[6]</sup>分别为源语言文档和目标语言文档估计一个基于主题的相关性模型,在此基础上,Wang等<sup>[7]</sup>提出了基于主题模型和伪相关反馈的跨语言信息检索系统,取得了良好的实验结果;刘伟成等<sup>[8]</sup>提出基于核典型相关分析算法的中英跨语言信息检索方法,实验结果证明了该方法的有效性。

基于查询扩展的跨语言信息检索研究的核心问题是如何得到原查询相关的扩展词实现跨语言查询扩展,以改善和提高跨语言信息检索性能。其典型的工作有微软亚洲研究院 Gao 等<sup>[9]</sup>提出的两步伪相关反馈法,在此基础上,吴丹等<sup>[10]</sup>通过计算前列  $n$  篇初检文档的语词权重并降序排列,抽取权重比较高的前列  $m$  个语词作为扩展词实现跨语言查询扩展,同时深入研究跨语言查询翻译前扩展、翻译后扩展以及两者结合的扩展效果。Lam 等<sup>[11]</sup>结合词分布表示方法和查询词与扩展词间的互信息计算扩展词权值实现跨语言信息检索,黄名选<sup>[12]</sup>提出一种基于矩阵加权关联模式的印尼-中跨语言信息检索模型,均取得了良好的实验效果。

当前,针对东盟小语种越南语言的跨语言信息检索研究报道不是很多,而中国与越南的政治、经济、文化等往来更加频繁和密切,面向越南语言的跨语言信息检索研究显得更加迫切和重要。为此,在原有的研究基础上,文中开展面向东盟国家语言的跨语言信息检索研究,提出一种基于完全加权关联规则后件扩展的越英跨语言信息检索模型。

## 2 融合完全加权模式挖掘与查询扩展的越英跨语言信息检索模型

### 2.1 模型基本思想

检索模型将越南语查询式通过机器翻译系统翻译为英文查询式,在原始英文文档集中检索英文文档,提取前列  $n$  篇初检英文文档提交给用户,通过用户的相关性判断得到用户相关反馈初检文档,采用基于支持度-条件概率增量比-兴趣度评价框架的完全加权词间关联规则挖掘技术对该文档进行挖掘,得到与原查询相关的关联规则模式,从规则模式中提取扩展词,和原查询组合成新查询再次检索原始英文文档集,得到最终检索结果,即英文文档结果集,再通过机器翻译系统将结果文档翻译为越南语文档,返回给用户。

### 2.2 模型结构图及其功能模块

根据上述检索模型的基本思想,给出了融合完全加权模式挖掘与查询扩展的越英跨语言信息检索模型结构,如图1所示。该检索模型结构由6个模块和2个数据库组成,即文本翻译模块、文本检索模块、初检结果用户相关反馈模块、面向越英跨语言检索的完全加权关联模式挖掘模块、越英跨语言查询扩展词生成模块和越英跨语言查询扩展实现模块,以及完全加权关联规则库和英文扩展词库等。各个模块的功能简述如下:

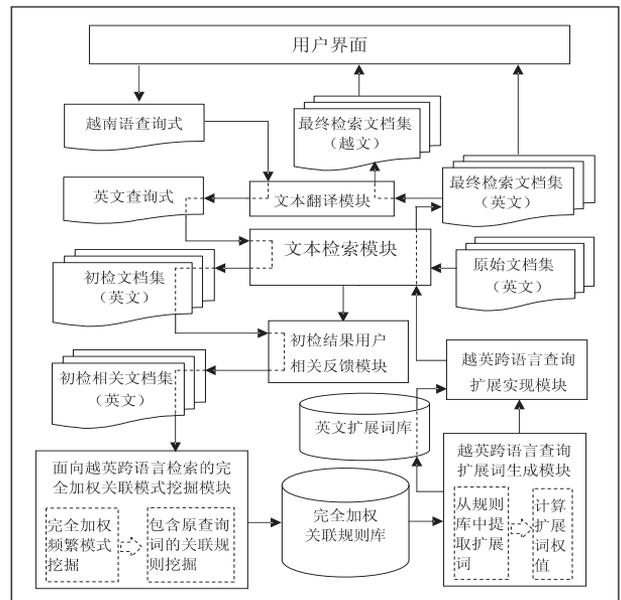


图1 融合完全加权模式挖掘与查询扩展的越英跨语言信息检索模型结构

**文本翻译模块:**该模块的功能是将越南语用户查询式翻译为英文查询式,将最终检索结果英文文档翻译为越南语文档。文本采用现有的机器翻译系统作为文本翻译模块,即微软的必应机器翻译接口(Microsoft translator API)。

**文本检索模块:**该模块负责将翻译后的查询和原始文档集中的文档进行相似度计算,并降序排列相似度,将排在前列的相似文档返回给用户。文中采用经典的、传统的向量空间模型方法计算查询与文档的相似度,即计算查询向量和文档向量之间夹角的余弦值作为其相似度值。

**初检结果用户相关反馈模块:**该模块负责收集用户对初检英文前列文档的相关性判断结果,提取被判断为相关的初检文档构建用户相关反馈文档集,同时,对该文档集进行英文词干提取、去除停用词和提取特征词等预处理。

**面向越英跨语言检索的完全加权关联模式挖掘模块:**该模块的功能是运用基于支持度-条件概率增量比-兴趣度评价框架的完全加权词间关联规则挖掘技

术对用户相关反馈文档集进行挖掘,得到包含有原查询词项的完全加权频繁项集和词间关联规则模式,构建完全加权关联规则库。

越英跨语言查询扩展词生成模块:该模块负责从完全加权关联规则库中提取扩展词,计算其权值,构建英文扩展词库。

越英跨语言查询扩展实现模块:该模块负责从英文扩展词库提取扩展词,将扩展词和原查询词组合为新查询提交到文本检索模块对原始文档集再次进行检索,得到最终检索的英文文档集,并将其送到文本翻译模块翻译为越南语文档集,最后将最终检索结果的越南语文档及其对应的英文文档返回给用户界面。

### 2.3 模型关键技术

#### 2.3.1 面向越英跨语言检索的完全加权词间关联模式挖掘

面向越英跨语言检索的完全加权词间关联模式挖掘基本思想是:采用基于支持度-条件概率增量比 CPIR-兴趣度评价框架的完全加权关联模式挖掘技术,结合用户原查询信息,对越英跨语言检索初检结果用户相关反馈文档集挖掘与原查询相关的完全加权频繁项集和关联规则模式。其中采用文献[13]的支持度计算公式计算关联模式的支持度,将条件概率增量比 CPIR 作为关联模型的置信度,其计算公式详见文献[14],对关联模式进行兴趣度评估,其评估公式采用文献[15]的关联模式兴趣度计算公式。在挖掘过程中,采取如下剪枝策略:剪除项集频次或者权值为 0 的候选项集;当挖掘到 2\_项集时,只保留含有原查询词的候选 2\_项集,其他的剪除;从候选 2\_项集以后,对应每个候选项集,如果包含该项集的后续 (k+1)\_项集权值阈值 KIWT(k, k+1)<sup>[16]</sup> 大于该项集权值 W(k),根据文献[16]的定理,可以剪除该候选项集。

上述挖掘思想可以形式化为算法 AWAPM\_CLIR (all-weighted association patterns mining for cross language information retrieval):

算法 1:AWAPM\_CLIR。

输入:Q<sub>v</sub>(越南语用户查询),Doc<sub>CLIR</sub>(跨语言初检英文相关反馈文档集),ms(最小支持度阈值),mc(最小置信度阈值)和 mi(最小兴趣度阈值);

输出:awAR<sub>q</sub>(含有译后英文原查询词项的完全加权关联规则模式集合)。

Begin

1. Let awFI<sub>q</sub> ← ∅; awAR<sub>q</sub> ← ∅;

//清空完全加权频繁项集模式集合 awFI<sub>q</sub> 及其关联规则模式集合 awAR<sub>q</sub>

2. Preproc (Doc<sub>CLIR</sub>);

//对初检英文相关文档进行词干提取、去除停用词和提取英文特征词等预处理工作。模型中所用的英文词干提取

程序是 Porter 程序(见 <http://tartarus.org/~martin/Porter-Stemmer>)

3. MiningawL<sub>1</sub>(Doc<sub>CLIR</sub>);

//挖掘完全加权频繁 1\_项集 awL<sub>1</sub>

3.1 扫描文档集 Doc<sub>CLIR</sub>,计算完全加权候选 1\_项集 C<sub>1</sub> 的权值 w(C<sub>1</sub>)

3.2 计算文档集 Doc<sub>CLIR</sub>中所有项目权值的总和 W

3.3 if (w(C<sub>1</sub>)/W ≥ ms) then L<sub>1</sub> ← C<sub>1</sub>, awFI<sub>q</sub> ← awFI<sub>q</sub> ∪ L<sub>1</sub>

3.4 计算 KIWT(1,2)的值。KIWT(1,2)<sup>[16]</sup>的计算公式见文献[16]

4. for (k=2; C<sub>k</sub> ≠ ∅; k++)

//挖掘含有原查询词的完全加权频繁 k\_项集 L<sub>k</sub>

4.1 if (W(C<sub>k-1</sub>) < KIWT(k-1, k)) then {剪除该候选项集 C<sub>k-1</sub>}

4.2 C<sub>k</sub> ← CreateC<sub>k</sub>(C<sub>k-1</sub>); //通过 A priori 连接<sup>[17]</sup>产生候选项集 C<sub>k</sub>

4.3 if (k=2) then {剪除不含英文查询词的候选 2\_项集 C<sub>2</sub>}

4.4 计算 C<sub>k</sub> 权值 w(C<sub>k</sub>)和 KIWT(k, k+1)的值。KIWT(k, k+1)<sup>[16]</sup>的计算公式见文献[16]

4.5 if (w(C<sub>k</sub>)=0) then {剪除该候选项集 C<sub>k</sub>}

4.6 if (w(C<sub>k</sub>)/W ≥ ms) then L<sub>k</sub> ← C<sub>k</sub>, awFI<sub>q</sub> ← awFI<sub>q</sub> ∪ L<sub>k</sub>

4.7 if (k 大于给定的所挖掘的候选项集长度阈值) then break;

5. 对于完全加权频繁项集模式集合 awFI<sub>q</sub>中的每个频繁项集 L<sub>k</sub>,挖掘中所有含原查询词的完全加权关联规则(q → Et),其中,q 为 L<sub>k</sub>中的原查询项项目,Et 为 L<sub>k</sub>中除了原查询项项目以外的项目

5.1 计算规则 q → Et 的条件概率增量比 CPIR 的值及其兴趣度的值。CPIR 值的计算公式见文献[14],规则兴趣度的计算公式见文献[15]

5.2 if (CPIR 的值 ≥ mc and 兴趣度 ≥ mi) then awAR<sub>q</sub> ← awAR<sub>q</sub> ∪ {q → Et};

6. output(awAR<sub>q</sub>);

//输出含有查询词的完全加权关联规则

End

#### 2.3.2 越英跨语言扩展词生成

文中跨语言扩展词来源于关联规则 q → Et 中的后件 Et 词项,并将其条件概率增量比 CPIR 和兴趣度作为扩展词权值的计算依据。因此,给出如下扩展词权值 W<sub>Et</sub> 的计算公式:

$$W_{Et} = \max(\text{CPIR 值} + \text{规则兴趣度值}) \quad (1)$$

此外,对于原查询项的权值  $W_q$  的计算详见文献[18]。

根据上述的越英跨语言扩展词产生的基本思想,给出生成算法 GetCLawET。

算法2: GetCLawET。

输入:  $awAR_q$  (完全加权关联规则模式集合);

输出:  $En\_ExpTerm$  (英文扩展词)。

Begin

1. for 对于  $awAR_q$  集合中的每个规则  $q \rightarrow Et$  do

{

提取规则  $q \rightarrow Et$  的后件  $Et$ ;

根据式1计算扩展词的权值;

$En\_ExpTerm \leftarrow En\_ExpTerm \cup Et$ , 并将扩展词及其权值

存入英文扩展词库;

}

2. EOutput( $En\_ExpTerm$ );

//输出英文扩展词

End;

### 2.3.3 融合完全加权模式挖掘与查询扩展的越英跨语言检索算法

根据上述给出的基本思想,文中提出的越英跨语言检索算法(Vietnamese-English cross language information retrieval based on all-weighted patterns mining and query expansion, VECLIR\_AWAM&QE)具体的描述如下:

算法3: VECLIR\_AWAM&QE。

输入:  $Q_v$  (越南语用户查询),  $ms, mc, mi, n$  (跨语言初检前列文档数);

输出: 越英跨语言检索结果(即越南语文档和英文文档)。

Begin

1.  $Q_E \leftarrow Q_v \text{ MachineTran}Q_v(Q_v)$ ;

//接受越南语用户查询式,将其翻译为英文查询式  $Q_E$

2.  $Doc_{CLR} \leftarrow Q_E \text{ Text Retrieval\_First}(Q_E, n)$ ;

//初检英文文档,构建用户相关反馈文档集  $Doc_{CLR}$

for 原始英文文档集中的每一篇文档  $Doc_i$  do

{

计算查询  $Q_E$  与文档  $Doc_i$  的相似度并降序排列;

提取前列  $n$  篇初检文档提交给用户进行相关性判断;

根据用户的相关性判断结果,构建初检用户相关反馈文档

集  $Doc_{CLR}$

}

3.  $awAR_q \leftarrow AWAPM\_CLR(Q_v, ms, mc, mi)$ ;

//调用上述挖掘算法 AWAPM\_CLR,运行后得到含有原查询项的完全加权关联规则模式集合  $awAR_q$

4.  $En\_ExpTerm \leftarrow GetCLawET(awAR_q)$ ;

//调用算法 GetCLawET,得到英文扩展词及其权值

5.  $NewQ_E \leftarrow Q_E \cup En\_ExpTerm$ ;

//将原查询和扩展词组合为新查询

6.  $En\_Doc \leftarrow NewQ_E \text{ Text Retrieval}(NewQ_E)$ ;

//扩展后的新查询再次检索得到最终英文文档  $En\_Doc$

6.1 for 原始英文文档集中的每一篇文档  $Doc_i$  do

计算新查询  $NewQ_E$  与文档  $Doc_i$  的相似度并降序排列;

6.2 提取前列文档作为最终检索结果英文文档  $En\_Doc$ ;

7.  $Viet\_Doc \leftarrow D_v \text{ MachineTran}D_v(En\_Doc)$ ;

//通过机器翻译系统将最终检索结果英文文档翻译为越南语文档  $Viet\_Doc$

8. VEDocOutput( $Viet\_Doc, En\_Doc$ );

//输出扩展后的检索结果:英文文档和越南语文档

End

## 3 实验设计及结果分析

### 3.1 实验数据及实验硬件环境

实验数据是 NTCIR-5 CLIR 提供的英文语料,即 Korea Times 的 2001 年新闻文本(简称为 ktn01),共 25.3 MB,以及 Mainichi Daily News 的 2000 年和 2001 年新闻文本(简称为 mdn00 和 mdn01),共 9.9 MB,共计 26 224 篇,其中 ktn01 为 14 069 篇,mdn00 为 6 608,mdn01 为 5 547。采用 TITLE 查询和 DESC 查询进行实验。NTCIR-5 CLIR 语料没有越南语版本的查询集,邀请专门翻译机构将 NTCIR-5 CLIR 中文版 50 个查询主题人工翻译为越南语查询,作为实验过程中用到的越南语用户查询。

### 3.2 基准实验及评价指标说明

文中实验对比算法是单语言检索算法(monolingual retrieval baseline, MRB)、越英跨语言检索算法(cross-language retrieval baseline, CLRB)和传统的基于伪相关反馈扩展的越英跨语言检索算法(Vietnamese-English cross-language retrieval based on pseudo relevance feedback, VECLR\_PRF),即 MRB 指用英文查询直接检索英文文档得到的检索结果,CLR B 指将越南语查询经机器翻译为英文后检索英文文档得到的检索结果,VECLR\_PRF 指按照文献[10]的跨语言扩展方法来实现越英跨语言信息检索。具体实验参数是:跨语言初检前列伪相关文档数是 20 篇,跨语言扩展词数是其权值较高的前列 20 个特征词。使用 R-查准率(R-prec)作为实验评价指标。

### 3.3 实验结果及其分析

为了便于实验的开展,将跨语言初检前列  $n$  篇文档中包含的已知结果集中的相关文档当作用户进行的相关反馈信息结果,构建用户相关反馈文档集。运行 MRB、CLR B、VECLR\_PRF 和 VECLIR\_AWAM&QE 等源程序,提交 50 个越南语的 TITLE 和 DESC 查询进行基准检索实验,得到各个算法的实验结果  $R\_prec$  值如表 1 和表 2 所示。实验参数设置如下:  $n = 50$ , 挖掘的项集长度为 2,支持度  $ms$  变化时:  $ms \in [0.001, 0.002, 0.003, 0.004, 0.005]$ ,  $mc = 0.01$ ,  $mi = 0.0001$ , 置信度

mc 变化:  $mc \in [0.01, 0.04, 0.06, 0.08, 0.1]$ ,  $ms = 0.005$ ,  $mi \in [0.001, 0.003, 0.005, 0.007, 0.009]$ 。  
 $0.001$ ,  $mi = 0.0001$ , 兴趣度  $mi$  变化:  $mc = 0.01$ ,  $ms =$

表 1 文中算法与基准对比算法的检索性能(R<sub>prec</sub>)比较(Title 查询)

检索算法	参数	Relax			Rigid		
		mdn00	mdn01	ktn01	mdn00	mdn01	ktn01
MRB		0.306 1	0.437 8	0.239 8	0.270 3	0.354 0	0.174 8
CLRB		0.281 8	0.282 9	0.220 7	0.240 0	0.237 7	0.184 0
VECLR_PRF		0.296 9	0.229 9	0.092 0	0.220 4	0.182 0	0.067 7
	ms	0.552 0	0.540 8	0.376 2	0.461 1	0.415 0	0.318 7
VECLR_AWAM&QE	mc	0.553 3	0.547 8	0.326 3	0.461 0	0.415 2	0.285 6
	mi	0.452 8	0.443 8	0.340 2	0.396 2	0.341 8	0.297 3

表 2 文中算法与基准对比算法的检索性能(R<sub>prec</sub>)比较(Desc 查询)

检索算法	参数	Relax			Rigid		
		mdn00	mdn01	ktn01	mdn00	mdn01	ktn01
MRB		0.240 4	0.326 9	0.160 0	0.207 3	0.349 8	0.144 5
CLRB		0.265 8	0.342 5	0.203 3	0.228 8	0.317 1	0.176 1
VECLR_PRF		0.198 8	0.252 7	0.090 8	0.168 2	0.278 3	0.075 5
	ms	0.486 9	0.536 4	0.379 5	0.404 1	0.477 3	0.326 2
VECLR_AWAM&QE	mc	0.465 1	0.551 8	0.368 7	0.387 2	0.495 2	0.327 6
	mi	0.433 0	0.484 6	0.321 9	0.383 7	0.449 0	0.289 3

由表 1 和表 2 可以看出,在 TITLE 和 DESC 查询实验中,在支持度阈值 ms、置信度阈值 mc 和兴趣度阈值 mi 等参数分别变化的情况下,算法 VECLR\_AWAM&QE 的检索结果的 R<sub>prec</sub> 值均比 MRB、CLRB 和 VECLR\_PRF 的高,效果比较显著。结果表明提出的越英跨语言信息检索模型是有效的,能有效地减少查询主题漂移,改善和提高跨语言检索性能。

## 4 结束语

将完全加权关联模式挖掘技术引入越英跨语言信息检索模型,结合用户相关反馈,提出一种基于完全加权关联规则后件扩展的越英跨语言信息检索模型,阐述了检索模型结构及其各个功能模块,以及实现的关键技术,进行了深入的实验分析,实验结果证明了该模型的有效性。下一步是将该越英跨语言检索模型实用化,以便应用到实际的检索系统中。

## 参考文献:

- [1] GERALDO A, MOREIRA V P. UFRGS@CLEF2008: using association rules for cross-language information retrieval[C]// Proceedings of the 9th cross-language evaluation forum conference on evaluating systems for multilingual and multimodal information access. Aarhus, Denmark: Springer-Verlag, 2008: 66-74.
- [2] 姚寒冰,王丽清,徐永跃. 供需信息跨语言检索算法研究[J]. 计算机技术与发展, 2017, 27(8): 152-155.
- [3] 吴丹,何大庆,王惠临. 一种基于相关反馈的跨语言信息检索查询翻译优化技术研究[J]. 情报学报, 2012, 31(4): 398-406.
- [4] GIANG L T, HUNG V T, PHAP H C. Experiments with query translation and re-ranking methods in Vietnamese-English bilingual information retrieval[C]// Proceedings of the fourth symposium on information and communication technology. New York, NY, USA: Association for Computing Machinery, 2013: 118-122.
- [5] GIANG L T, HUNG V T, PHAP H C. Building structured query in target language for Vietnamese-English cross language information retrieval systems[J]. International Journal of Engineering & Technical Research, 2015, 4(4): 146-151.
- [6] DEBASIS G, JOHANNES L, GARETH J F J. Cross-lingual topical relevance models[C]// Proceedings of the 24th international conference on computational linguistics. Mumbai, India: [s. n.], 2012: 927-942.
- [7] WANG Xuwen, WANG Xiaojie, ZHANG Qiang. A web-based CLIR system with cross-lingual topical pseudo relevance feedback[C]// International conference of the cross-language evaluation forum for European languages. Berlin: Springer, 2013: 104-107.
- [8] 刘伟成,张志清,孙吉红. 基于 KCCA 的跨语言专利信息检索研究[J]. 情报科学, 2010, 28(5): 751-755.
- [9] GAO Jianfeng, NIE Jianyun, ZHANG Jian, et al. TREC-9

- Areas in Communications, 2012, 30(3):497-508.
- [4] 黄俊伟,周朋光,张仁迟,等.超密集网络中小小区分簇和子载波分配算法[J].电子技术应用,2017,43(7):104-109.
- [5] GE Xiaohu, TU Song, MAO Guoqiang, et al. 5G ultra-dense cellular networks[J]. IEEE Wireless Communications, 2016, 23(1):72-79.
- [6] 王俊才,刘婷婷,杨晨阳,等.超密集网络中最大化网络吞吐量的预测资源分配[J].信号处理,2017,33(3):260-267.
- [7] GOLREZAEI N, MOLISCH A F, DIMAKIS A G, et al. Femtocaching and device-to-device collaboration: a new architecture for wireless video distribution[J]. IEEE Communications Magazine, 2013, 51(4):142-149.
- [8] 牛腾,张冬梅,许魁,等.基于D2D的无线多媒体网络编码广播重传策略[J].中国科学:信息科学,2018,48(2):205-220.
- [9] 秦恒加,米志超,董超,等.Android多跳D2D通信系统的实现与优化[J].计算机科学,2017,44(2):135-139.
- [10] 邱永芳,邱恭安,周永箐.基于可变距离的车联网D2D通信连接选择[J].电信科学,2017,33(5):39-45.
- [11] 谭泽辛.基于交互实时流的无线D2D调度策略优化研究[D].长沙:湖南大学,2015.
- [12] WU Xinzhou, TAVILDAR S, SHAKKOTTAI S, et al. FlashLinQ: a synchronous distributed scheduler for peer-to-peer ad hoc networks[C]//48th annual Allerton conference on communication, control, and computing. Allerton, IL, USA: IEEE, 2010:514-521.
- [13] NADERIALIZADEH N, AVESTIMEHR A S. ITLinQ: a new approach for spectrum sharing in device-to-device communication systems[J]. IEEE Journal on Selected Areas in Communications, 2013, 32(6):1139-1151.
- [14] KIM Y, BACCELLI F, DE VECIANA G. Spatial reuse and fairness of ad hoc networks with channel-aware CSMA protocols[J]. IEEE Transactions on Information Theory, 2014, 60(7):4139-4157.
- [15] GOEBBELS S, JENNEN R. Enhancements in wireless broadband networks using smart caching an analytical evaluation [C]//International symposium on personal, indoor and mobile radio communications. Cannes, France: IEEE, 2008:1-5.
- [16] DEGHAN M, SEETHARAM A, JIANG B, et al. On the complexity of optimal routing and content caching in heterogeneous networks[C]//Computer communications. [s. l.]: IEEE, 2014:936-944.
- [17] GOLREZAEI N, MOLISCH A F, DIMAKIS A G. Base-station assisted device-to-device communications for high-throughput wireless video networks[C]//IEEE international conference on communications. Ottawa, ON, Canada: IEEE, 2012:7077-7081.
- [18] JI M, TULINO A M, LLORCA J, et al. Order-optimal rate of caching and coded multicasting with random demands [J]. IEEE Transactions on Information Theory, 2017, 63(6):3923-3949.
- [19] GOLREZAEI N, SHANMUGAM K, DIMAKIS A G, et al. FemtoCaching: wireless video content delivery through distributed caching helpers [C]//Proceedings of INFOCOM. Orlando, FL, USA: IEEE, 2013:1107-1115.
- [20] PARK S H, SIMEONE O, SAHIN O, et al. Robust and efficient distributed compression for cloud radio access networks [J]. IEEE Transactions on Vehicular Technology, 2013, 62(2):692-703.
- [21] KIM J, CAIRE G, MOLISCH A F. Quality-Aware Streaming and Scheduling for Device-to-Device Video Delivery[J]. IEEE/ACM Transactions on Networking, 2016, 24(4):2319-2331.

(上接第 168 页)

- CLIR experiments at MSRCN[C]//Proceedings of the 9th text retrieval evaluation conference. [s. l.]: [s. n.], 2001:343-353.
- [10] 吴丹,何大庆,王惠临.基于伪相关反馈的跨语言查询扩展[J].情报学报,2010,29(2):232-239.
- [11] GIANG L T, HUNG V T, PHAP H C. Improve cross language information retrieval with pseudo-relevance feedback [J]. International Journal of Engineering Research & Technology, 2015, 4(6):1-7.
- [12] 黄名选.基于矩阵加权关联模式的印尼中跨语言信息检索模型[J].数据分析与知识发现,2017,1(1):26-36.
- [13] 周秀梅,黄名选.基于项权值变化的矩阵加权关联规则挖掘[J].计算机应用研究,2015,32(10):2918-2923.
- [14] WU X D, ZHANG C Q, ZHANG S C. Efficient mining of both positive and negative association rules[J]. ACM Transactions on Information Systems, 2004, 22(3):381-405.
- [15] 周秀梅,黄名选.基于项权值变化的完全加权正负关联规则挖掘[J].电子学报,2015,43(8):1545-1554.
- [16] 黄名选,严小卫,张师超.基于矩阵加权关联规则挖掘的伪相关反馈查询扩展[J].软件学报,2009,20(7):1854-1865.
- [17] AGRAWAL R, IMIELINSKI T, SWAMI A. Mining association rules between sets of items in large database[C]//Proceedings of 1993 ACM SIGMOD international conference on management of data. Washington D C, USA: ACM, 1993:207-216.
- [18] SALTON G, BUCKLEY C. Term-weighting approaches in automatic text retrieval[J]. Information Processing & Management, 1988, 24(5):513-523.