段落检索及其相关算法研究

付鸿鹄[1,2] 张晓林[1]

[1](中国科学院文献情报中心 北京 100080) [2](中国科学院研究生院 北京 100049)

【摘要】 总结段落检索及其涉及的段落划分和相关算法,讨论文本分割和段落抽取的差别,介绍并比较几种常用的段落划分方法以及几类段落检索算法,在此基础上对段落检索的研究方向进行展望。

【关键词】 段落检索 信息检索 文本分割 【分类号】 TP31

Research on Passage Retrieval and the Algorithms

Fu Honghu[1,2] **Zhang Xiaolin**[1]

[1](*Library of Chinese Academy of Sciences, Beijing 100080, China*)

[2](*Graduate School of the Chinese Academy of Sciences, Beijing 100049, China*)

【**Abstract**】 This paper provided an overview of passage retrieval, including approaches of division of documents and retrieval algorithms. The difference between text segmentation and passage extraction is also discussed. Several kinds of passage retrieval algorithms and several approach of division of documents are introduced. After comparing of these algorithms and approaches, some research directions of passage retrieval are presented.

【**Keywords**】 Passage retrieval Information retrieval Text segmentation

随着网络技术的迅速发展和普及,用户已经可以快速便捷地检索出大量文档,这些文档中往往只有部分段落内容与用户需求密切相关,但大量地浏览被检索的文档往往妨碍用户快速筛选出相关的段落。研究表明[1],当文档很长或者包含多个主题段落时,以段落为检索对象、针对用户查询提供更加精确的文档段落(而不是提供整篇文档),能有效提高用户检索效率,这也使得段落检索成为信息检索的一种重要形式。

1 段落检索

段落检索(Passage Retrieval)算法是 Salton 等在 1993 年首次提出的,旨在发现一个文档中内容相似的段落,即采用某种策略以确认两个片段彼此相似[2]。在实际检索中,往往将用户检索式作为一个"对照段落",通过检索找出与用户检索式足够相似的段落以及所在文档。

段落检索有助于显著提高用户的检索效率[2]:

(1)提高文档相关性确定的准确性:许多情况下用户检索式各个检索词个别地散落在文档的多个段落中,仅按检索词出现次数而提供的文档可能实际上是不相关的。而段落检索按照检索词在一个段落中出现次数来确定与用户检索式的"相关性",避免散布的检索词造成的误检,避免文档长度对相关度的影响,检索精度得到提高。

(2)提高相关内容呈现的有效性:当通过检索发现"相关段落"后,可直接将"确实相关的内容(段落)"呈现出来,而不是将"可能包含相关内容的整个文档"显示给用户。而且,可以按照段落相关性来排列显示段落本身,可以按照逻辑关系显示相互关联的多个段落。

然而,段落检索也存在一定的困难,不仅段落检索要求有可靠、健壮的方法从文档中分离出段落,而且文档划分成段落意味着会产生更多的需要排序的项,导致检索有更高的计算复杂度。另外,对于段落是变长的段落划分方案,长度归一化的问题仍然存在。更为重要的是,以段落排序为基础的全文检索,可能会忽略那些没有高相似度段落而文档总体相关度较高的相关文档。

2 段落检索中的段落划分

2.1 段落划分的基本类别

根据段落划分与段落相关度计算的时间关系,段落划分分为文本分割和段落抽取。

文本分割(Text Segmentation,TS)指的是首先把每个文档划分成段落,然后再判断各个段落是否与用户检索式相关。通过文本分割产生的文档段落可以是重叠的,即一个段落的结束点可能出现在其后面一个段落的开始

收稿日期:2006-12-27

点之后；文本分割可以保留文本结构，这样生成的段落（可能是文档的章、节）就有可能以通过层次机构关联起来。段落提取（Passage Extraction，PE）指的是从整篇文档出识别出与用户检索式相关的段落，不必事先把文档划分成静态的片段。相关段落可能会比一个自然段或章节大，也可能会小或者重叠，即提取的段落和文档的本身的段落不一定相匹配。

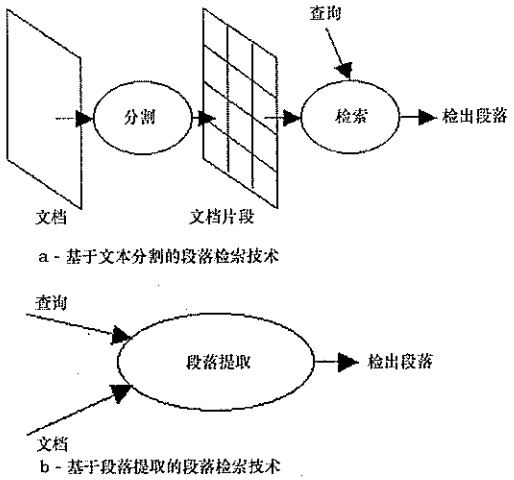

图 1 文本分割和段落提取[3]

图 1 展示了文本分割和段落提取两种段落检索技术的差别。文本分割可以在进行检索前就将文档划分为静态的段落，一次段落划分的结果能够尽可能的满足多次不同的检索需求；段落提取在检索过程中按检索比较结果动态产生段落，从而依据每个用户的具体需求来提供最好、最适用的段落。

2.2 段落划分的具体方法

无论是文本分割还是段落抽取，合理的段落划分是有效的段落检索的关键。在以往的研究中，主要采用有四种方式来实现段落的划分。

（1）按照章节结构划分[1,4]

使用文档的逻辑结构如章节（Sections）来划分段落。即以文本中的标记信息作为段落的划分依据，如：章节（Sections）、自然段（Paragraphs）、句子（Sentences）等。以文档逻辑结构作为段落划分的依据，保留了文档的逻辑结构信息和语义特征，段落是相对完整的信息单元，保证了信息的相对完整性。但文档逻辑结构的可辨认性往往是问题，而且由于逻辑上的段落存在长度上的差异，在处理时必须进行长度的归一化。

（2）按照等长自然段划分[5]

通过粗糙的等长的方式把文档分成不相连的片段。这种方法是为了避免按章节划分出现的长度差异较大和结构难以识别等问题，通过把相邻的自然段（Paragraphs）聚集在一起来确定段落，使每一个分块大小大于等于预定义的字节数，同时保证自然段的完整，即分段的边界与自然段边界一致。等长的段落划分避免了长度归一化的问题，但相邻自然段的聚集以长度为依据而不是文档的逻辑结构和语义信息，一定程度上存在段落和逻辑信息单元的不匹配的问题。

（3）按照话题迁移划分[6-8]

段落总是关于某个话题，不同的段落往往是关于不同的话题，因此可以根据由自动检测的话题（Topic）的迁移来确定段落。技术上以句子或自然段为单位，通过对比相邻句子或自然段的相似度、并将相似度低于一定限定判断为话题迁移，从而切分段落。这种方法一定程度上保证了段落的逻辑和语义，但系统处理上有比较高的复杂度。

（4）按照固定长度词序列划分[9]

按照固定长度词序列划分段落又称为“窗口段落”，直接按照相等的词汇量来切分段落，第一个段落的起始位置可以由文档中出现的第一个检索词的位置确定，后续每个段落切分的起点可以是上个段落的末点或中点（有利于避免切分所造成的逻辑内容的分离）。这种划分方法操作简单，也避免了长度归一化的问题，但它完全忽略了文档的逻辑结构和语义特征，划分效率受到影响。

在现有条件下，由于大量的文档并不提供计算机可读的逻辑结构上的标识，计算机很难识别出文档的章节，按章节结构来划分文档段落存在较大的困难；按照话题迁移来划分段落需要进行大量的比较和计算，实现复杂度较高；而按等长自然段和固定长度词序列的划分方法，实现复杂度较低。研究表明[2]，等长的段落划分由于段落划分简单易行，同时避免的长度归一化的问题，有着较低的计算复杂度，能够获得最高的处理效率。

3 段落检索的相关算法

目前，用于计算段落与检索提问相关度的段落检索算法主要有如下几种类型。

3.1 基于词频统计的算法

词频统计算法是通过计算文档段落中与检索提问相匹配的词语出现的频率判断段落的相关度。这类算法最容易实现也易于理解，但由于只是通过词频统计来计算，检索精确度并不高。

MITRE 算法[10]采用“词重叠算法”（Word Overlap Algorithm），计算在一个句子和检索提问相同的关键词（即重叠词）的数量。早期的算法只简单计算重叠词数量，后来为了提高精确度，通过加权对算法进行了改进，引入了重叠集（Overlap Set）和最大重叠集（MaxOset（Maximal Overlap Set））的概念。定义如下：

$O_{w,q} = \{s \mid s \cap q = w\}$

Ω_q = q 的全部重叠集

maximal($O_{w,q}$)如果$\forall O_{v,q} \in \Omega_q$,$w \not\subset v$其中,

$M_q = \{O_{w,q} \in \Omega_q \mid \text{maximal}(O_{w,q})\}$

$C_q = \{s \mid s$ 符合查询 $q\}$

q 表示一个提问,可以看作是一个词的集合;s 表示一个句子,也可以看作是一个词的集合;w、v 表示重叠词集,即每个句子中与提问中相同的词的集合。重叠集 $O_{w,q}$定义为一个句子的集合,该集合中的每个句子与检索提问 q 的重叠词集 w 都是相同的,即与提问 q 包含相同的词集 w 的句子的集合;Ω_q 表示对于提问 q 的全部重叠集。最大重叠集 maximal($O_{w,q}$)是一个特殊的一个重叠集,该重叠集中的重叠词集 w 不是任何其他重叠词集 v 的子集;M_q 表示对于提问 q 的最大重叠集的集合。

例如:

对于自然语言的提问:How much was Babe Belanger paid to play amateur basketball?

有如下 5 个句子:

S1: She was a member of the winningest basketball team Canada ever had.

S2: Babe Belanger never made a cent for her skills.

S3: They were just a group of young women from the same school who liked to play amateur basketball.

S4: Babe Belanger played with the Grads from 1929 to 1937.

S5: Babe never talked about her fabulous career.

重叠集包括:({S1},{S2,S4},{S3},{S5}),其中最大重叠集为:({S2,S4},{S3})。

C_q 为候选结果集,其中包含符合提问 q 的句子。这些句子一般来自于最大重叠集,由于包含有比较多的关键词,往往具有比较高的相关度。

3.2 基于检索词密度的算法

为了提高检索的精确度,许多基于检索词密度的算法被应用到段落检索中。这类算法通过计算段落中出现的检索词之间的距离来进行段落评分。由于考虑了段落中出现的多个检索词之间的距离关系,与词频统计相比精确度有了一定的提高,但其算法杂度也相应地提高。同时,这种距离关系还只是一种位置上的关系,并不能完全揭示检索词之间在逻辑和语义上的相互关系。IBM[11]、SiteQ[12]、MultiText[13]等采用的就是这种算法。

(1)IBM[11]

IBM 的算法通过计算一系列的距离值来进行段落相关度计算和排序。

首先把每个段落分成句子;围绕每个句子形成一个窗口,每个窗口大小为 3 个句子;计算 5 个距离值:词匹配度(Matching Words)、辞典匹配度(Thesaurus Match)、未匹配词度(Mis-match Words)、离散度(Dispersion)和词聚合度(Cluster Word)。这些度经过加权计算,线性地组合在一起最后给出段落相关度的最终评分。其中,5 个距离值的定义和在评分中的作用如下:

词匹配度为出现在检索式和段落中的词的 tf/idf 值之和(+);

辞典匹配度为出现在检索词的 WordNet 同义词出现在段落中的词的 tf/idf 值之和(+);

未匹配词度为出现在查询中但不出现在段落中的词的 tf/idf 值(-);

离散度为两个匹配查询术语之间的出现的词数(-);

词聚合度为在提问和段落中都相临出现的词数(+)。

(2)SiteQ[12]

SiteQ 的算法定义了三种形式的关键词:词条形式(Lemma form)、词干形式(Stemmed form)和 WordNet 词义形式。词条形式通过它在文本中的位置加权,一个专有名词或者一个大写字母开头的普通名词和最高级形容词将比动词、形容词和副词的权重要高。词干形式和它的词条形式享有同样的权重。WordNet 词义匹配的关键词权重最低,和它的构成词数相关。对于多个句子的段落,计算每个单独句子的评分之和,作为段落的评分。

$$Score_{sent} = Score_1 + Score_2$$

$$Score_1 = \sum_i wgt(qw_i) \text{(如果 } qw_i \text{ 在句中出现)}$$

$$Score_2 = \frac{\sum_{j=1}^{k-1} \frac{wgt(dw_j) + wgt(dw_{j+1})}{\alpha \times dist(j,j+1)^2}}{k-1} \times matched\_cnt$$

其中:

$wgt(qw_i)$:查询词 i 的权值

$wgt(dw_j)$:与文档中词 j 相匹配的查询词 i 的权值

$dist(j,j+1)$:文档中词 j 和 j+1 的距离

matched_cnt:句子中与查询词匹配的词数

α:常数

即,每个句子按检索式密度进行评分:每个句子的项和检索式进行匹配获得 Score1,计算匹配的项之间的距离获得 Score2,句子的得分 Scoresent 为 Score1 和 Score2 之和。每一个检索式按词条形式、WordNet 词义和词干形式依次与文档项进行匹配尝试,并取第一次匹配的项的类型的权值。

(3)MultiText[13]

MultiText 的段落检索算法对于包含有许多查询词的短段落给与高的评分。在这个算法中段落窗口(Passage Window)以查询词为开始和结束,评分基于段落中出现的查询词的次数和窗口的大小。

每个文档被作为一个词的序列来处理:$D = d_1 d_2 d_3 \cdots d_m$

在每一个词的位置(1…m),有多个索引项来标引这个词的信息。这些索引项包含了这个词本身、它的词干和一些表示这个词的位置对应于一个名字、数字或其他一些值的符号。

查询被作为一个项集(a set of terms)来处理:$Q = \{q_1, q_2, q_3, \cdots\}$

其中,每个项(term)是一个词、短语或词干。

定义段落 extent(u,v),$1 \le u \le v \le m$ 为文档 D 的一个子序列,它从位置 u 开始,到位置 v 结束。即:extent(u,v) = $d_u d_{u+1} d_{u+2} \cdots d_v$。项集 T 为 Q 的一个子集($T \subseteq Q$),如果 extent(u,v)包含了 T 中的全部项,则称 extent(u,v)满足 ;如果 extent(u,v)满足 T,并且不包含一个也满足 T 的自序列,即不存在一个子序列 extent(u',v'),$u < u' \le v' < v$ 或 $u < v' \le u' < v$ 满足 T,则称 extent(u,v)是 T 的一个覆盖(Cover)。系统计算出文档集中所有文档包含的对 Q 的所有子集的覆盖。

每个段落获得的分值由其长度和所匹配的项的权值确定。查询项 t 的权值定义为 $w_t = \log(N/f_t)$,其中 f_t 是 t 在整个文集中出现的次数,N 是整个文集中全部文档的长度之和。

项集 $T \subseteq Q$ 的权值为 中包含的所有项的权值之和:

$$W(T) = \sum_{t \in T} w_t$$

如果段落 extent(u,v)是 T 的一个覆盖,则它的得分由其长度和它所匹配的项的权值确定:$C(T,u,v) = W(T) - |T| \log(v-u+1)$。当得分较高的段落选出后,分别计算其中点位置$(u+v)/2$,并以此为中心截取 200 个词作为命中段落返回。

3.3 基于语言模型的算法

1998 年,Ponte 和 Croft 首先将语言模型应用到文档检索中,认为文档的相关性可以通过文档"产生"查询的可能性来衡量[14]。在信息检索环境中下,语言建模的问题就成为一个在给定一个文档的语言模型情况下(有或没有查询的语言模型),估计文档和查询由同一语言模型生成的可能性的问题。2002 年,Liu 和 Croft 将语言模型与段落检索结合起来,提出基于语言模型的段落检索方法[15]。新加坡国立大学 Cui H. 等人提出的模糊关系匹配(Fuzzy Relation Matching)的方法[16]也是一类基于语言模型的段落检索算法。

模糊关系匹配算法[16]是 Cui H. 等人在开放领域问答系统(QA)的研究中为了提高段落检索效率来改善问答系统性能而提出的一种基于语言模型的段落检索算法。该算法通过构建文档和查询提问的语言模型,计算相邻词之间的依赖关系来进行段落评分。系统以句子为单位划分段落,把每个句子作为一个段落来处理。

首先借用 Minipar[17](一个高效、健壮的依赖关系解析器)生成句子依赖树,从中抽取词间的关系路径。在一个依赖树中,每个结点代表一个词或者短语,并用链接(link)表示从这个结点(主结点,Governor)到其修饰结点(Modifier Node)的指向关系。在两个结点间定义一个关系路径。从提问和句子中分别抽取出成对关系路径后,根据从提问中抽取的路径计算从句子中抽取的路径的匹配值。如图 2 为从提问 < Question > 和句子 < S1 > 抽取出来的关系路径。

< Question > What percent of the nation's cheese does Wisconsin produce?

< S1 > In Wisconsin, where farmers produce roughly 28 percent of the nation's cheese, the outrage is palpable.

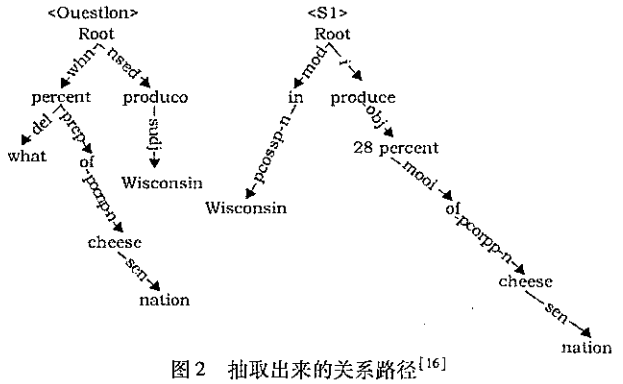

图 2 抽取出来的关系路径[16]

通过扩展 IBM 统计转换模型[18]来计算路径间的匹配值。从提问 Q 和句子 S 中标示出两条对应的路径,分别表示为 PQ 和 PS,长度分别记为 m 和 n。转换概率 Prob(PS|PQ)为所有的排列之和。

$$\mathrm{Prob}(P_S|P_Q) = \frac{\zeta}{m^n} \sum_{\alpha_1=1}^{m} \cdots \sum_{\alpha_n=1}^{m} \prod_{i=1}^{n} P_t(\mathrm{Rel}_i^{(S)} | \mathrm{Rel}_{\alpha_i}^{(Q)}) \quad (1)$$

其中,$\mathrm{Rel}_i^{(S)}$ 表示路径 P_S 中的第 i 个关系,$\mathrm{Rel}_{\alpha_i}^{(Q)}$ 表示路径 P_Q 中对应的关系。关系的排列通过 α_i 给出,α_i 表示给定关系 $\mathrm{Rel}_i^{(S)}$ 在提问中对应的关系。ξ 是一个小的常数。$P_t(\mathrm{Rel}_i^{(S)} | \mathrm{Rel}_{\alpha_i}^{(Q)})$ 表示关系转换概率。系统假设每一条关系都能够被转换,当 $\mathrm{Rel}_i^{(S)}$ 和 $\mathrm{Rel}_{\alpha_i}^{(Q)}$ 相同时,$P_t(\mathrm{Rel}_i^{(S)} | \mathrm{Rel}_{\alpha_i}^{(Q)})$ 为 1(当一个关系转换成自身时,转换概率最大)。只考虑最可能的排列,基于关系转换概率,通过为每一个句子路径中的关系找到其在问题路径中的最可能的映射来计算最可能的排列,这样,路经转换概率简化成:

$$\mathrm{Prob}(P_S|P_Q) = \frac{\xi}{m^n} \prod_{i=1}^{n} P_t(\mathrm{Rel}_i^{(S)} | \mathrm{Rel}_{A_i}^{(Q)}) \quad (2)$$

其中,A_i 表示最可能的排列。此外,可以只使用 P_S 的长度 n 来规范化等式(2)。通过转换,去除常数,P_S 的匹配值:

$$\mathrm{MatchScore}(P_S) = \mathrm{Prob}(P_S|P_Q) = \frac{\xi}{n} \sum_{i=1}^{n} \mathrm{Log}P_t(\mathrm{Rel}_i^{(S)} | \mathrm{Rel}_{A_i}^{(Q)}) \quad (3)$$

最后,计算每个路径的匹配值之和作为句子与提问的关系匹配值。这个值反映了句子和提问的匹配程度:高的得分表示句子中出现的提问中的词具有和提问中相同的语义,具有高的相关度。

4 结 语

(1)不同的段落划分方法各有其优缺点,在实际应用中,往往需要根据具体情况,如文本集的特点、所选用的文档检索系统、段落检索算法等来选择。基于等长的段落划分方法虽然能取得较高的处理效率,但由于损失了文档逻辑结构信息,并不能满足特定的检索需求。在实际应用中,段落划分方法

能够达到的处理效果与检索的目标也有密切关系,需要根据不同的目标选择不同的段落划分方法。随着对检索结果逻辑和语义上的关注越来越多,基于文档逻辑结构和语义结构的段落划分方法也越来越多地被研究和采用。

(2)不同的算法在相关度计算的处理上各具特色,计算复杂度和处理效果也各有差异,适应不同的应用需求。如,基于词频统计的算法原理简单易懂,实现起来复杂度较低,但精确度不高;基于检索词密度的算法考虑检索词的距离关系,检索精确度有了一定的提高,在各项研究中较多地被采用,如 MultiText 就是在 TREC[19] 的问答系统研究中采用最多的算法之一;基于语言模型的算法考虑检索词间的语义关系,对于提高检索精确度有一定的作用,但相应的也增加了计算复杂度。此外,对部分算法在问答系统中的应用效果进行的测试表明[20],在采用不同的文档检索系统时,段落检索算法也显示出不同的检索效果。现有信息环境中,信息资源的极大丰富使用户对信息检索的知识化的要求越来越高,相应的,段落检索也必将朝着知识化的方向发展,基于语言模型的检索算法、基于文档的语义关系[21]提高检索效率和质量等相关研究将进一步发展。

(3)段落检索与许多研究领域密切结合,如在文本分类、问答系统、搜索引擎技术等研究中都有相关的应用。以往的研究表明使用文档段落作为信息的基本单位,计算一个文档和提问之间的相关性,能够明显地提高信息检索系统的效果。知识化和智能化是信息检索的主要方向,基于段落检索,分析段落间的逻辑和语义关系,实现知识化检索是段落检索一个应用研究方向。

参考文献:

1 Salton G, Allan J, Buckley C. Approaches to Passage Retrieval in Full Text Information Systems. In: Proceedings of the 16th Annual International ACM SIGIR Conference. Pittsburgh, PA, 1993: 49 - 58

2 Kaszkeil M, Zobel J. Passage Retrieval Revisited, In: Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '97). 1997: 178 - 185

3 Melucci M. Passage Retrieval: A Probabilistic Technique . Information Processing & Management, 1998, 34(1): 43 - 68

4 Wilkinson R. Effective Retrieval of Structured Documents. In: Proc. ACM - SIGIR International Conference on Research and Development in Information Retrieval. Dublin, Ireland, 1994:311 - 317

5 Zobel J, Moffat A, Wilkinson R, et al. Efficient Retrieval of Partial Documents. Information Processing & Management, 1995, 31(3): 361 - 377

6 Hearst M A, Plaunt C. Subtopic Structuring for Full - Length Document Access. In Proc. ACM - SIGIR International Conference on Research and Development in Information Retrieval. Pittsburg, 1993: 59 - 68

7 Knaus D, Mittendorf E, Schauble P, et al. Highlighting Relevant Passages for Users of the Interactive SPIDER Retrieval System. In: NIST Special Publication 500 - 236: Text Retrieval Conference (TREC - 4), Washington, 1995:233 - 243

8 Mittendorf E, Schauble P. Document and Passage Retrieval Based on Hidden Markov Models. In Proc. ACM - SIGIR International Conference on Research and Development in Information Retrieval, Dublin, Ireland, 1994: 318 - 327

9 Callan J P. Passage - level Evidence in Document Retrieval. In Proc. ACM - SIGIR International Conference on Research and Development in information Retrieval, Dublin, Ireland, 1994: 302 - 309

10 Light M, Mann G S, et al. Analyses for Elucidating Current Question Answering Technology. Journal of Natural Language Engineering, Special Issue on Question Answering, 2001

11 Ittycheriah A, Franz M, et al. IBM's Statistical Question Answering System. In Proceedings of the 9th Text Retrieval Conference (TREC - 9), 2000:229

12 Lee G G, Seo J, et al. SiteQ: Engineering High Performance QA System Using Lexico - semantic Pattern Matching and Shallow NLP. In: Proceedings of the Tenth Text REtrieval Conference (TREC 2001), 2001:442

13 Clarke C L A, Cormack G V, et al. Question Answering by Passage Selection. In: Proceedings of the 9th Text Retrieval Conference (TREC - 9), 2000:673

14 Ponte J M, Croft W B. A Language Modeling Approach to Information Retrieval. In: Proceedings of the 21st Annual international ACM SIGIR Conference on Research and Development in information Retrieval (SIGIR'98), 1998:275 - 281

15 Liu X, Croft W B. Passage Retrieval Based on Language Models, In: Proceedings of the 11th International Conference on Information and Knowledge Management(CIKM'02), 2002, 375 - 382

16 Cui H, Sun R, et al. Question Answering Passage Retrieval Using Dependency Relations. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '05), 2005:400 - 407

17 Lin D. Dependency - based Evaluation of MINIPAR. In: Proc. of Workshop on the Evaluation of Parsing Systems, Granada, Spain, 1998

18 Al - Onaizan, Curin J Y, et al. Statistical Machine Translation. Final Report, JHU Workshop. 1999

19 Text Retrieval Conference. http://trec. nist. gov/(Accessed Jan 12, 2007)

20 Tellex S, Katz B, Lin J, et al. Quantitative Evaluation of Passage Retrieval Algorithms for Question Answering, SIGIR'03, July 28 - August 1, 2003, Toronto, Canada

21 Ofoghi B, Yearwood J, Ghosh R. A Semantic Approach to Boost Passage Retrieval Effectiveness for Question Answering. In: Proceedings of the 48th Conference on Computer Science (CRPITS'48), 2006

(作者 E - mail:fuhh@mail.las.ac.cn)